THEORY OF COMPUTING, Volume 9 (24), 2013, pp. 759–781
www.theoryofcomputing.org

SPECIAL ISSUE: APPROX-RANDOM 2012

Hardness of Vertex Deletion and Project Scheduling*

Ola Svensson†

Received November 4, 2012; Revised May 4, 2013; Published September 26, 2013

Abstract: Assuming the Unique Games Conjecture, we show strong inapproximability results for two natural vertex-deletion problems on directed graphs: for any integer $k \geq 2$ and arbitrary small $\varepsilon > 0$, the Feedback Vertex Set problem and the DAG Vertex Deletion problem are inapproximable within a factor $k - \varepsilon$ even on graphs where the vertices can be almost partitioned into k solutions. This gives a more structured and yet simpler UG-hardness result for the Feedback Vertex Set problem than the previous hardness result (albeit using the "It Ain't Over Till It's Over" Theorem).

In comparison with the classical Feedback Vertex Set problem, the DAG Vertex Deletion problem has received little attention. Although we think it is a natural and interesting problem, the main motivation for our inapproximability result stems from its relationship with the classical Discrete Time–Cost Tradeoff Problem. More specifically, our results imply that the deadline version is UG-hard to approximate within any constant. This explains the difficulty in obtaining good approximation algorithms for that problem and further motivates previous alternative approaches such as bicriteria approximations.

ACM Classification: F.1.3, G.1.6, G.2.2

AMS Classification: 68Q17, 68W25, 68Q25

Key words and phrases: inapproximability, graphs, scheduling, vertex-deletion problems

*An extended abstract of this paper appeared in the Proceedings of the 15th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems, 2012 [19].

†This research was supported by Grant228021-ECCSciEng of the European Research Council.

DOI: 10.4086/toc.2013.v009a024

1 Introduction

Many interesting discrete optimization problems can be formulated as the problem of finding, in a given (directed) graph, a large induced subgraph with a desired property. One of the most studied such problems is the *Feedback Vertex Set (FVS)* problem where the property is acyclicity, i. e., given a directed graph $G = (V,E)$ we wish to delete the minimum number of vertices so that the resulting graph is acyclic. Another example is the *DAG Vertex Deletion (DVD)* problem, where we are given an integer k and a directed acyclic graph and we wish to delete the minimum number of vertices so that the resulting graph has no path of length[1] k.

The FVS problem and the related Feedback Arc Set problem (FAS) were shown to be NP-complete already in Karp's seminal paper [9] and there is a long history of approximation algorithms for these problems. Leighton and Rao [13] first gave a $O(\log^2 |V|)$-approximation algorithm for FAS. Seymour [17] gave an approximation guarantee of $O(\log |V| \log\log |V|)$ for FVS by showing that a certain linear program approximates the value within a factor $O(\log |V| \log\log |V|)$. Seymour's arguments were then generalized by Even et al. [5] to obtain the best known approximation algorithms achieving a factor $O(\log |V| \log\log |V|)$ for both problems on weighted graphs. For weighted instances, easy reductions show that FVS and FAS are in fact equivalent in terms of approximability [5].

Motivated by certain VLSI design and communication problems, Paik et al. [15] considered the DVD problem and showed it to be NP-complete on general graphs and polynomial time solvable on series-parallel graphs. One can also see that DVD for a fixed k is a special case of the Vertex Cover problem on k-uniform hypergraphs and has a fairly straightforward k-approximation algorithm.

In comparison to FVS, the DVD problem has received little attention. Our main motivation for studying its approximability comes from its relationship (that we prove in Section 5) with the classical deadline version of the project scheduling problem known as the *Discrete Time–Cost Tradeoff problem* (DTCT). Informally (see Section 5 for a formal definition), this is the problem where we are given a deadline and a project consisting of tasks related by precedence constraints, and the time it takes to execute each task depends, via a given cost function, on how much we pay for it. The objective is to minimize the cost of executing all the tasks in compliance with the precedence constraints so that they all finish within the given deadline. Due to its obvious practical relevance, the problem has been studied in various contexts over the last 50 years (see the paper [11] by Kelly and Walker for an early reference). Fulkerson [6] and Kelley [10] obtained polynomial time algorithms if all cost functions are linear. In contrast, the problem becomes NP-hard for arbitrary cost functions [3] and there is even no known constant-factor approximation algorithm in the general case. However, better (approximation) algorithms have been obtained for special cases. For example, Grigoriev and Woeginger [7] gave polynomial time algorithms for special classes of precedence constraints and one of several algorithms by Skutella [18] is a bicriteria approximation that, for any $\mu \in (0,1)$, approximates DTCT within a factor $1/(1-\mu)$ if the deadline is allowed to be violated by a factor $1/\mu$.

In summary, there are no known constant-factor approximation algorithms for FVS, DVD, and DTCT although few strong inapproximability results are known. The best known NP-hardness of approximation results follow from the fact that they are all as hard to approximate as Vertex Cover which is NP-hard to

[1]For notational convenience, we shall measure the length of a path as the number of vertices it contains instead of the number of edges.

approximate within a factor 1.3606 [4]. It is indeed easy to see that Vertex Cover is a special case of FVS and DVD, and Grigoriev and Woeginger [7] gave an approximation-preserving reduction from Vertex Cover to DTCT. If we assume the Unique Games Conjecture (UGC) [12], our understanding of the approximability of FVS becomes significantly better: the hardness of approximation result for Maximum Acyclic Subgraph by Guruswami et al. [8] implies that it is NP-hard to approximate FVS within any constant factor assuming the UGC. However, the results in [8] use very sophisticated techniques that are not known to imply a similar hardness for DVD and DTCT. Specifically, the techniques in [8] are inspired by Raghavendra's [16] beautiful results on constraint-satisfaction problems that relate UG-hardness results to the integrality gap of a canonical semidefinite programming relaxation. To apply a similar framework, the authors of [8] treat ordering problems as n-ary constraint-satisfaction problems, show that one can restrict the domain to a large constant, and make a similar connection between integrality gaps of a semidefinite programming relaxation and UG-hardness. The framework developed in [8] is very powerful and leads to optimal UG-hardness results for all so-called ordering-constraint satisfaction problems. However, the hardness reductions obtained from integrality gaps in this way are not explicit in the sense that it is difficult to grasp if the instances obtained have any particular structure that can be useful for obtaining hardness results for other problems, such as DTCT and DVD.

Even though the initial motivation of this work was to better understand the approximability of DTCT (and DVD), the techniques that we develop also lead to a more structured UGC-based hardness result for FVS: similar to the recent results for Vertex Cover on k-uniform hypergraphs by Bansal and Khot [1, 2], we show that, for any integer $k \geq 2$ and arbitrarily small $\varepsilon > 0$, there is no $(k-\varepsilon)$-approximation algorithm for FVS even on graphs where the vertices can be almost partitioned into k feedback vertex sets. Our reduction is also much simpler and more "direct" than the one in [8] (albeit using the "It Ain't Over Till It's Over" Theorem) but is tailored to FVS and does not yield any inapproximability result for the Maximum Acyclic Subgraph problem. As already noted, feedback vertex set and feedback arc set are equivalent for weighted instances [5] and therefore a non-constant UG-hardness result for weighted feedback arc set also follows from our results. (The hardness result in [8] is also for weighted instances.) More importantly, our techniques also lead to an analogous result for the DVD problem (and thereby for DTCT). Formally, our results for the vertex-deletion problems considered can be stated as follows.

Theorem 1.1. *For any integer $k \geq 2$ and arbitrary constant $\varepsilon > 0$, the following problems are* UG*-hard:*

FVS: *Given a graph $G(V,E)$, distinguish between the following cases:*

- (Completeness): *there exist k disjoint subsets $V_1,\dots,V_k \subset V$ satisfying $|V_i| \geq \frac{1-\varepsilon}{k}|V|$ and such that a subgraph induced by any $k-1$ of these subsets is acyclic.*
- (Soundness): *every induced subgraph of $\varepsilon|V|$ vertices has a directed cycle.*

DVD: *Given a DAG $G(V,E)$, distinguish between the following cases:*

- (Completeness): *there exist k disjoint subsets $V_1,\dots,V_k \subset V$ satisfying $|V_i| \geq \frac{1-\varepsilon}{k}|V|$ and such that a subgraph induced by any $k-1$ of these subsets has no path of length k.*
- (Soundness): *every induced subgraph of $\varepsilon|V|$ vertices has a path of length $|V|^{1-\varepsilon}$.*

Note that completeness says that if we let $V' = V \setminus (V_1 \cup \cdots \cup V_k)$ then the sets $V' \cup V_i$ for $i = 1, \ldots, k$ are almost disjoint solutions of size at most $(1/k + \varepsilon)|V|$ each. In contrast, soundness says that any solution basically needs to delete all vertices (even to avoid paths of length $|V|^{1-\varepsilon}$ for DVD).

When proving UGC-based inapproximability results, the main task is usually to design "gadgets" for the problems considered that simulate a so-called dictatorship test. Once we have such "dictatorship gadgets," the process of obtaining UGC-based hardness results often follows from (by now) fairly standard arguments. In particular, the main ideas needed for our reductions leading to Theorem 1.1 are already present in the design of the gadgets. We have therefore chosen to present those gadget constructions with less cumbersome notation in Section 3 and the reductions from Unique Games can be found in Section 4.

As alluded to above, our main interest in DVD stems from its relationship with DTCT. More specifically, in Section 5, we give an approximation-preserving reduction from DVD to DTCT that combined with Theorem 1.1 yields:

Theorem 1.2. *It is* UG*-hard to approximate DTCT within any constant factor.*

This explains the difficulty in obtaining good approximation algorithms for DTCT and also further motivates alternative approaches such as the bicriteria approach by Skutella [18] that approximates the DTCT within a constant factor if the deadline is allowed to be violated by a constant factor.

2 Preliminaries

2.1 Low-degree influence and "It Ain't Over Till It's Over" Theorem

Let $[k] = \{0, 1, \ldots, k-1\}$. When analyzing our hardness reductions, we shall use known properties regarding the behavior of functions of the form $f : [k]^R \to \{0, 1\}$ depending on whether they have influential coordinates. Similar to [14, Section 3], we define the influence of the i-th coordinate by

$$\mathrm{Infl}_i(f) = \mathbb{E}_x[\mathrm{Var}_{x_i}(f(x))]\,.$$

We note that if $f : \{-1, 1\}^R \to \{-1, 1\}$ then this definition coincides with the intuitive expression

$$\Pr_x[f(x_1, \ldots, x_i, \ldots, x_R) \neq f(x_1, \ldots, -x_i, \ldots, x_R)]\,.$$

It is well known that if we let $f = \sum_\Phi \hat{f}(\phi) X_\phi$ be the multi-linear representation of f (where, in analogy with the standard Fourier representation, the characters $(X_\phi)_{\phi \in [k]^R}$ define an orthonormal basis of the vector space of all functions $[k]^n \to \mathbb{R}$) then the influence can also be expressed as

$$\mathrm{Infl}_i(f) = \sum_{\phi : \phi_i \neq 0} \hat{f}^2(\phi)\,,$$

which motivates the following definition of the degree-d influence of the i-th coordinate:

$$\mathrm{Infl}_i^d(f) = \sum_{\phi : \phi_i \neq 0, |\phi| \leq d} \hat{f}^2(\phi)\,.$$

As we shall not work directly with these definitions or with the multi-linear representation, we refer the reader to [14] for the precise definitions and cut the discussion short by mentioning the property of low-degree influence that will be crucial to us. This property follows from the fact that $\sum_\phi \hat{f}^2(\phi) = \mathbb{E}_x[f(x)^2] \leq 1$.

Observation 2.1 (see, e. g., Proposition 3.8 in [14]). For a boolean function $f : [k]^R \to \{0,1\}$, the sum of all degree-d influences is at most d.

We shall now introduce a simplified version of the "It Ain't Over Till It's Over" Theorem that is sufficient for the applications in this paper. The first proof was given by Mossel et al. [14, Theorem 4.9] and a more combinatorial proof of a simplified version (very similar to the one used here) was given by Bansal and Khot [1, Theorem 3.1] who used it to prove tight inapproximability results for Vertex Cover and a classical single-machine scheduling problem. In fact many of our ideas are inspired by [1]. For $x \in [k]^R$ and a subset $S_\varepsilon \subseteq R$ of εR indices, let

$$C_{x,S_\varepsilon} = \{z \in [k]^R : z_j = x_j \ \forall j \notin S_\varepsilon\}$$

denote the sub-cube defined by fixing the coordinates not in S_ε according to x. Let also $f(C_{x,S}) \equiv 0$ denote the expression that f is identical to 0 on the sub-cube C_{x,S_ε}.

Theorem 2.2 (special case of Theorem 4.9 in [14]). *For every $\varepsilon, \delta > 0$ and integer k, there exists $\eta > 0$ and integer d such that any $f : [k]^R \to \{0,1\}$ that satisfies*

$$\mathbb{E}[f] \geq \delta \qquad \textit{and} \qquad \forall i \in [R], \mathrm{Infl}_i^d(f) \leq \eta\,,$$

has

$$\Pr_{x,S_\varepsilon}[f(C_{x,S_\varepsilon}) \equiv 0] \leq \delta\,.$$

Here and throughout the paper, the probability over x, S_ε is such that x and S_ε are taken independently uniformly at random. When ε is clear from the context we often also abbreviate S_ε by S. We remark that, as the above statement only considers whether a function is identical to 0 on a chosen sub-cube, the randomly picked set S_ε can be interpreted as a truncated version of the operator V in [14] that picks the set by including each vertex independently with probability ε' (by selecting $\varepsilon' \ll \varepsilon$ sufficiently small only an insignificant fraction of the outcomes will be disregarded). This slight change is not necessary but it will simplify the explanation of the reductions in this paper.

The theorem says that a reasonably balanced function with no low-degree influential coordinates has very low probability to be identical to 0 over the random choice of sub-cubes. In contrast, it is easy to see that a dictatorship function (on the boolean domain) $f(x) = x_s$, for some s, has

$$\Pr_{x,S_\varepsilon}[f(C_{x,S_\varepsilon}) \equiv 0] = \Pr_{x,S_\varepsilon}[f(C_{x,S_\varepsilon}) \equiv 1] \geq 1/2 - \varepsilon\,.$$

It is this drastic difference that we will exploit in our hardness reductions.

2.2 Unique Games Conjecture

An instance of Unique Games $\mathcal{L} = (G(V,W,E),[R],\{\pi_{v,w}\}_{(v,w)})$ consists of a regular bipartite graph $G(V,W,E)$ and a set $[R]$ of labels. For each edge $(v,w) \in E$ there is a constraint specified by a permutation $\pi_{v,w} : [R] \to [R]$. The goal is to find a labeling $\rho : (V \cup W) \to [R]$ so as to maximize $\mathrm{val}(\rho) := \Pr_{e \in E}[\rho \text{ satisfies } e]$, where a labeling ρ is said to satisfy an edge $e = (v,w)$ if $\rho(v) = \pi_{v,w}(\rho(w))$. For a Unique Games instance $\mathcal{L}$, we let

$$\mathrm{OPT}(\mathcal{L}) = \max_{\rho : V\cup W \to [R]} \mathrm{val}(\rho)\,.$$

The now famous Unique Games Conjecture that has been extensively used to prove strong hardness of approximation results can be stated as follows.

Conjecture 2.3 ([12])**.** *For any constants $\zeta,\gamma > 0$, there is a sufficiently large integer $R = R(\zeta,\gamma)$ such that, for Unique Games instances $\mathcal{L}$ with label set $[R]$ it is NP-hard to distinguish between:*

- (Completeness): $\mathrm{OPT}(\mathcal{L}) \geq 1 - \zeta$.
- (Soundness): $\mathrm{OPT}(\mathcal{L}) \leq \gamma$.

3 Dictatorship gadgets for vertex-deletion problems

We give simple gadgets for the vertex-deletion problems considered that informally correspond to a dictatorship test in the following sense: (Completeness:) any dictatorship function $f : [k]^R \to [k]$ (defined by $f(x) = x_s$ for some $s \in [R]$) corresponds to a good solution whereas (Soundness:) any non-trivial solution corresponds to a function $f : [k]^R \to \{0,1\}$ with a high-influence coordinate. These gadgets are then used in Section 4 together with a more general form of the "It Ain t Over Till It's Over" Theorem (see Theorem 4.1) to obtain analogous hardness results assuming the Unique Games Conjecture. Specifically, Theorem 4.1 generalizes Theorem 2.2 in the sense that it tests whether at least two of a given set of functions have a common coordinate with high influence instead of only verifying if a single function has a high-influence coordinate. This generalization will allow us to check consistency of labels in the Unique Games instance in a fairly standard way. Apart from this difference, the reductions in Section 4 are similar to the gadget constructions. However, as Section 4 has a more cumbersome notation, we believe that the gadget constructions here present the underlying ideas more clearly.

Throughout this section, we fix k to be an integer, $\varepsilon, \delta > 0$ to be arbitrarily small constants, and let η and d be as in Theorem 2.2 (depending on k, ε and δ).

3.1 Feedback Vertex Set

Here we shall describe a graph $G = (V,E)$ that naturally corresponds to a dictatorship test in the following sense:

- *(Completeness:)* A dictatorship function partitions the vertex set into subsets $V', V_0, \ldots, V_{k-1}$ satisfying $V_j \geq \frac{1-\varepsilon}{k}|V|, |V'| \leq \varepsilon |V|$, and for $j \in [k]$ the graph obtained by deleting $V' \cup V_j$ is acyclic.

- *(Soundness:)* Any feedback vertex set that deletes less than $(1-2\delta)|V|$ vertices corresponds to a function $f:[k]^R \to \{0,1\}$ with a coordinate i so that $\mathrm{Infl}_i^d(f) > \eta$.

3.1.1 Dictatorship gadget

To make the analysis more intuitive, it will be convenient to first present a gadget that consists of two types of vertices that we refer to as *bit-vertices* and *test-vertices* and all arcs are between bit- and test-vertices:

- There is a bit-vertex b_x of weight ∞ for every $x \in [k]^R$.
- There is a test-vertex $t_{x,S}$ of weight 1 for every $x \in [k]^R$ and every subset $S \subseteq [R]$ of εR indices.
- The arcs incident to a test-vertex $t_{x,S}$ are the following. There is an arc $(b_z, t_{x,S})$ if $z \in C_{x,S}$ and an arc $(t_{x,S}, b_z)$ if $z \in C^{\oplus}_{x,S}$, where
$$C^{\oplus}_{x,S} = \{z \oplus 1 : z \in C_{x,S}\}$$
(here $\oplus$ denotes coordinatewise-addition mod k and 1 is the all-ones vector of appropriate dimension).

As the bit-vertices have weight ∞, they will never be deleted in an optimal solution. We can therefore obtain an unweighted graph G of the same optimal value by omitting the bit-vertices and having an arc $(t_{x,S}, t_{x',S'})$ between two test vertices if there exists a bit-vertex b_z so that $(t_{x,S}, b_z)$ and $(b_z, t_{x',S'})$. The vertex set of G will therefore correspond to the set T of test-vertices. The analysis of G therefore follows from proving that (completeness:) any dictatorship function partitions the test-vertices as required (Section 3.1.2) and (soundness:) that any solution that deletes less than a $(1-2\delta)$ fraction of the test-vertices corresponds to a function with a coordinate of high influence (Section 3.1.3).

3.1.2 Completeness

We show that a dictatorship function $f:[k]^R \to [k]$ of index s, i. e., $f(x) = x_s$, naturally partitions the test-vertices into subsets $T', T_0, \ldots, T_{k-1}$ satisfying $|T_j| \geq ((1-\varepsilon)/k)|T|, |T'| \leq \varepsilon |T|$, and such that the sets $T' \cup T_j$ for $j \in [k]$ are almost disjoint feedback vertex sets of size at most $(1/k+\varepsilon)|T|$ each.

As $f(x) = x_s$, it partitions the bit-vertices in k equal sized sets
$$B_j = \{b_x : f(x) = j\} \qquad \text{for} \qquad j \in [k]\,.$$

We say that a test-vertex $t_{x,S}$ is good if $s \notin S$ and partition the good test-vertices into k equal sized sets
$$T_j = \{t_{x,S} : s \notin S \text{ and } f(x) = j\} \qquad \text{for} \qquad j \in [k]\,.$$

The sets are of equal size since they are partitioned according to x and whether a test-vertex is good only depends on S. Furthermore, as at least a $(1-\varepsilon)$ fraction of the test-vertices are good we have that $|T_j| \geq ((1-\varepsilon)/k)|T|$ for $j \in [k]$ and therefore the remaining test-vertices in T' are at most $\varepsilon|T|$ many.

It remains to show that $T_j \cup T'$ defines a feedback vertex set for any $j \in [k]$. The key observation is that T_j only have incoming edges from bit-vertices in B_j and outgoing edges to bit-vertices in $B_{j\oplus 1}$. Indeed, consider a test-vertex $t_{x,S} \in T_j$ and an arc $(b_z, t_{x,S})$. By definition we have that $z \in C_{x,S}$ and as S is

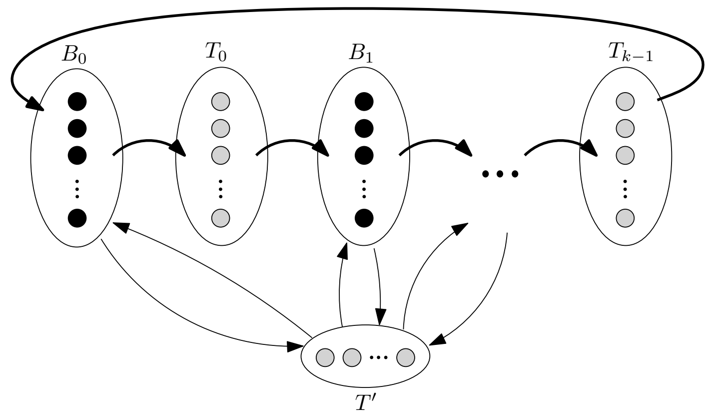

Figure 1: The structure of the graph after partitioning the vertices according to a dictatorship function. They gray and black vertices depict test-vertices and bit-vertices, respectively.

good we have that $f(z) = f(x) = j$, which shows that $z \in B_j$. We infer that $t_{x,S}$ only has outgoing edges to $B_{j\oplus 1}$ by the exact same argument.

By the above arguments we have that, in the graph without bad test-vertices, any cycle must visit at least one vertex in each of the sets $T_0, T_1, \ldots, T_{k-1}$ (see also Figure 1). This is because all outgoing edges from a test-vertex in T_j go to bit-vertices in $B_{j\oplus 1}$ and these bit-vertices have only outgoing edges to test-vertices in $T_{j\oplus 1}$ and so on. Therefore, the graph obtained by deleting all bad test-vertices and one of the sets $T_0, T_1, \ldots, T_{k-1}$ is acyclic as required.

3.1.3 Soundness

Let A be the last $1/2$ fraction of the bit-vertices according to a topological sort of the graph. Let f_A be the indicator function of A. Note that a test-vertex $t_{x,S}$ has incoming arcs from all bit-vertices in $C_{x,S}$ and outgoing arcs to all bit-vertices in $C^{\oplus}_{x,S}$. Therefore, if a test-vertex $t_{x,S}$ is not deleted then we must have that either f_A is identical to 0 on $C_{x,S}$ (if $t_{x,S}$ is placed before the last bit-vertex for which f_A evaluates to 0) or identical to 1 on $C^{\oplus}_{x,S}$ (if $t_{x,S}$ is placed after the last bit-vertex for which f_A evaluates to 0) depending on where $t_{x,S}$ is placed according to the topological sort.

As $\mathbb{E}[f_A] = 1/2$, we have by Theorem 2.2 that if $\mathrm{Infl}^d_i(f_A) \le \eta$ for all $i \in [R]$ then

$$\Pr_{x,S}[f(C_{x,S}) \equiv 0] \le \delta$$

and

$$\Pr_{x,S}[f(C^{\oplus}_{x,S}) \equiv 1] = \Pr_{x,S}[f(C_{x,S}) \equiv 1] = \Pr_{x,S}[(1-f)(C_{x,S}) \equiv 0] \le \delta\,.$$

Therefore, if the solution does not correspond to a function with a coordinate of high low-degree influence it must have deleted at least a $(1-2\delta)$ fraction of the test-vertices.

3.2 DAG Vertex Deletion problem

We shall describe a directed acyclic graph (DAG) $G = (V,E)$ that naturally corresponds to dictatorship test in the following sense:

- *(Completeness:)* A dictatorship function partitions the vertex set into subsets $V', V_0, \ldots, V_{k-1}$ satisfying $V_j \geq \frac{1-\varepsilon}{k}|V|, |V'| \leq \varepsilon|V|$, and such that for $j \in [k]$ the graph obtained by deleting $V' \cup V_j$ has no path of length k.
- *(Soundness:)* Any graph obtained by deleting less than $(1-6\delta)|V|$ vertices either has a path of length $|V|^{1-\delta}$ or corresponds to a function $f : [k]^R \to \{0,1\}$ with a coordinate i such that $\mathrm{Infl}_i^d(f) > \eta$.

3.2.1 Dictatorship gadget

As in Section 3.1, it will be convenient to first present a gadget that consists of two types of vertices that we refer to as *bit-vertices* and *test-vertices*, and all edges will be between bit- and test-vertices:

- The bit-vertices are partitioned into $L+1$ bit-layers (L is selected below). Each bit-layer $\ell = 0, \ldots, L$ contains a bit-vertex b_x^ℓ of weight ∞ for every $x \in [k]^R$.
- Similarly, the test-vertices are partitioned into L test-layers. Each test-layer $\ell = 0, \ldots, L-1$ has a test-vertex $t_{x,S}^\ell$ of weight 1 for every $x \in [k]^R$ and every subset $S \subseteq [R]$ of εR indices.
- The arcs are the following: there is an arc $(b_z^\ell, t_{x,S}^{\ell'})$ if $\ell \leq \ell'$ and $z \in C_{x,S}$, and there is an arc $(t_{x,S}^{\ell'}, b_z^\ell)$ if $\ell > \ell'$ and $z \in C_{x,S}^{\oplus}$.
- Finally, L is selected so as $\delta L \geq |T|^{1-\delta}$, where T is the set of test-vertices.

Note that, as there are only arcs from a bit-layer ℓ to a test-layer ℓ' if $\ell' \geq \ell$ and only arcs from a test-layer ℓ' to a bit-layer ℓ if $\ell > \ell'$, the constructed graph is acyclic. Similar to the gadget for FVS, the bit-vertices can be omitted to obtain an unweighted graph G (with the set T of test-vertices as vertices) with the same optimal value by having an arc between two test-vertices if there was a path between them through one bit-vertex. Note that a path in G of length k is a path in the gadget that consists of k test-vertices. When arguing about the gadget, we will therefore say that a *path has length k if it consists of k test-vertices*.

Similarly to Section 3.1, the analysis of G follows from proving that (completeness:) any dictatorship function partitions the test-vertices as required (Section 3.1.2) and (soundness:) that any solution that deletes less than a $(1-6\delta)$ fraction of the test-vertices either has a path of length $|T|^{1-\delta}$ or corresponds to a function with a coordinate of high influence (Section 3.1.3).

3.2.2 Completeness

We show that a dictatorship function $f : [k]^R \to [k]$ of index s naturally partitions the test-vertices into subsets $T', T_0, \ldots, T_{k-1}$ satisfying $T_j \geq ((1-\varepsilon)/k)|T|, |T'| \leq \varepsilon|T|$, and such that for $j \in [k]$ the graph obtained by deleting $T' \cup T_j$ has no path of length k.

This can be seen by the same arguments as in Section 3.1.2. Indeed if we "collapse" the different layers by identifying the different copies of bit- and test-vertices then the gadget constructed here is identical to the gadget in that section. We can therefore (by the arguments of Section 3.1.2), partition the bit-vertices into k equal sized sets $B_0, B_1, \dots, B_{k-1}$ and all but an ε fraction of the test-vertices into k equal sized sets $T_0, T_1, \dots, T_{k-1}$ so that any test-vertex in T_j has only incoming arcs from bit-vertices in B_j and outgoing arcs to bit-vertices in $B_{j\oplus 1}$.

Any $j \in [k]$ therefore corresponds to a solution by removing an ε fraction of the test-vertices (i. e., the set T') and those test-vertices in T_j.

3.2.3 Soundness

Before proceeding to the analysis it will be convenient to consider a different but equivalent formulation of the problem.

First, note that in any solution to DVD, i. e., a subgraph so that each path contains less than k test-vertices, we can find a coloring χ (using for example depth-first search) that assigns a color in $\{1,2,\dots,k\}$ to the bit-vertices with the property that, for each remaining test-vertex, the maximum color assigned to its predecessors is strictly less than the minimum color assigned to its successors. Similarly, any such coloring χ can be turned into a solution to DVD by deleting those test-vertices for which not all predecessors are assigned lower colors than all its successors. Furthermore, from the construction of the arcs, we can assume without loss of generality that the coloring satisfies $\chi(b_x^\ell) \le \chi(b_x^{\ell'})$ if $\ell \le \ell'$.

From the above discussion, the following equivalent formulation of DVD on the instances constructed can be obtained: find a coloring χ that assigns a color in $\{1,2,\dots,k\}$ to each bit-vertex satisfying $\chi(b_x^\ell) \le \chi(b_x^{\ell'})$ if $\ell \le \ell'$ so as to minimize the number of unsatisfied test-vertices where a test-vertex $t_{x,S}^\ell$ is said to be satisfied if

$$\max_{z\in C_{x,S}} \chi(b_z^\ell) < \min_{z\in C_{x,S}^{\oplus}} \chi(b_z^{\ell+1}),$$

that is, all its predecessors are assigned lower colors than any of its successors.

It will also be convenient to consider the following lower bound on the colors assigned to most bit-vertices in each layer: define the color $\chi(\ell)$ of a bit-layer $\ell = 0,1,\dots,L$ as the maximum color that satisfies $\Pr_x[\chi(b_x^\ell) \ge \chi(\ell)] \ge 1-\delta$.

Now, with each test-layer $\ell = 0,1,\dots,L-1$ we associate the indicator function $f^\ell : [k]^R \to \{0,1\}$ defined as follows

$$f^\ell(x) = \begin{cases} 0 & \text{if } \chi(b_x^{\ell+1}) > \chi(\ell), \\ 1 & \text{otherwise.} \end{cases}$$

The key observation for the soundness analysis is the following.

Claim 3.1. *For $\ell = 0,\dots,L-1$, assuming that $\mathrm{Infl}_i^d(f^\ell) \le \eta$ for all $i \in [R]$: if a 3δ fraction of the test-vertices of test-layer ℓ are satisfied, then $\chi(\ell) < \chi(\ell+1)$.*

Proof. As at least a 3δ fraction of the test-vertices of test-layer ℓ are satisfied,

$$\Pr_{x,S}\left[\max_{z\in C_{x,S}} \chi(b_z^\ell) < \min_{z\in C_{x,S}^{\oplus}} \chi(b_z^{\ell+1})\right] \ge 3\delta\,.$$

By the definition of $\chi(\ell)$ we have $\Pr_x[\chi(b_x^\ell) \ge \chi(\ell)] \ge 1-\delta$ and therefore

$$\Pr_{x,S}\left[\chi(\ell) < \min_{z\in C_{x,S}^{\oplus}} \chi(b_z^{\ell+1})\right] = \Pr_{x,S}\left[f^\ell(C_{x,S}) \equiv 0\right] \ge 2\delta\,.$$

As $\mathrm{Infl}_i^d(f^\ell) \le \eta$ for all $i \in [R]$, we conclude by using Theorem 2.2 that $E[f^\ell] < \delta$ and hence $\chi(\ell+1) > \chi(\ell)$. □

If a coloring satisfies more than a 6δ fraction of the test-vertices then at least a 3δ fraction of the test-layers are such that at least a 3δ fraction of the test-vertices of that layer are satisfied, which in turn by the preceding claim implies that either one of them corresponds to a function with a coordinate of high influence or $3\delta L$ many colors are needed (or equivalently the graph contains a path consisting of at least $3\delta L - 1 \ge \delta L \ge |T|^{1-\delta}$ test-vertices).

4 Hardness assuming the Unique Games Conjecture

In order to turn our dictatorship gadgets into hardness proofs (assuming the Unique Games Conjecture), we need a more general "It Ain't Over Till It's Over" Theorem that not only verifies that a given number of functions all are dictatorships but ideally they should also be dictators of the *same* coordinate. Again an even more general variant of the theorem follows from [14, Theorem 4.9] and an easier proof of a very similar version to the case presented here can be found in [1, Theorem 3.4].

Theorem 4.1 (special case of Theorem 4.9 in [14][2])**.** *For every $\varepsilon, \delta > 0$ and integer k, there exist $\eta > 0$ and integers t, d such that any collection of functions $f_1, \dots, f_t : [k]^R \to \{0,1\}$ that satisfies*

$$\forall j, \mathbb{E}[f_j] \ge \delta \qquad \textit{and} \qquad \forall i \in [R], \forall 1 \le \ell_1 \ne \ell_2 \le t, \min\left\{\mathrm{Infl}_i^d(f_{\ell_1}), \mathrm{Infl}_i^d(f_{\ell_2})\right\} \le \eta\,,$$

has

$$\Pr_{x,S_\varepsilon}\left[\bigwedge_{j=1}^{t} (f_j(C_{x,S_\varepsilon}) \equiv 0)\right] \le \delta\,.$$

For the applications of the above theorem that appear in this paper, the interesting conclusion can be reformulated as follows: if $\Pr_{x,S_\varepsilon}\left[\bigwedge_{j=1}^t f_j(C_{x,S_\varepsilon}) \equiv 0\right] > \delta$ for t fairly balanced functions then at least two of them must have a common influential coordinate.

In our (soundness) analyses, we associate a boolean function $f_w : [k]^R \to \{0,1\}$ with each $w \in W$ of the considered Unique Games instance. Let $N(v)$ denote the set of neighbors of a vertex v. We then

[2]To see that Theorem 4.1 is a special case of Theorem 4.9 in [14], consider the function $g : [k]^R \to [0,1]$ that outputs the average of the functions $f_1, f_2, \dots, f_t$, i. e., $g(x) = (1/t)\sum_{\ell=1}^t f_\ell(x)$. If the functions $f_1, f_2 \dots, f_t$ satisfy the premises then $\mathbb{E}[g] \ge \delta$ and g has no coordinate of high low-degree influence because $\mathrm{Infl}_i^d(g) \le (1/t)\mathrm{Infl}_i^d(f)$ and t is selected to be sufficiently large. Theorem 4.9 says that the probability that such a g is identical to 0 on a randomly chosen sub-cube is very small which is equivalent to our conclusion

$$\Pr_{x,S_\varepsilon}\left[\bigwedge_{j=1}^{t} \left(f_j(C_{x,S_\varepsilon}) \equiv 0\right)\right] \le \delta\,.$$

use the preceding theorem to test whether t (or more) neighbors $w_1, \ldots w_t \in N(v)$ of a vertex $v \in V$ are "close to" consistent, i. e., $f_{w_1}, \ldots, f_{w_t}$ are dictatorships on coordinates $\rho(w_1), \ldots, \rho(w_t)$ such that $\pi_{v,w_1}(\rho(w_1)) = \pi_{v,w_2}(\rho(w_2)) = \cdots = \pi_{v,w_t}(\rho(w_t))$. Indeed, if we let

$$f_{w_j} \circ \pi_{v,w_j}(x) = f_{w_j}(x_{\pi_{v,w_j}(1)}, x_{\pi_{v,w_j}(2)}, \ldots, x_{\pi_{v,w_j}(R)})$$

denote the function whose coordinates have been permuted according to π_{v,w_j} then on the one hand,

$$\Pr_{x,S_\varepsilon}\left[\bigwedge_{j=1}^{t}\left(f_{w_j} \circ \pi_{v,w_j}(C_{x,S_\varepsilon}) \equiv 0\right)\right] \geq 1/2 - \varepsilon \qquad \text{if they are consistent},$$

and on the other hand (assuming $f_{w_1}, \ldots, f_{w_t}$ are fairly balanced) if no two of them are "close to" consistent then Theorem 4.1 implies that

$$\Pr_{x,S_\varepsilon}\left[\bigwedge_{j=1}^{t}\left(f_{w_j} \circ \pi_{v,w_j}(C_{x,S_\varepsilon}) \equiv 0\right)\right] \leq \delta\,.$$

Similar to the gadget reductions, it is this drastic difference that we exploit to obtain our hardness results.

For the reductions, it shall be convenient to let $C_{x,S,v,w}$ denote the sub-cube

$$C_{x,S,v,w} = \{z : z_j = x_{\pi_{v,w}(j)}\ \forall j : \pi_{v,w}(j) \notin S\}\,,$$

i. e., the image of the sub-cube $C_{x,S}$ via $\pi_{v,w}$. Note that with this notation we have that

$$\Pr_{x,S}\left[\bigwedge_{j=1}^{t}\left(f_{w_j} \circ \pi_{v,w_j}(C_{x,S}) \equiv 0\right)\right] = \Pr_{x,S}\left[\bigwedge_{j=1}^{t}\left(f_{w_j}(C_{x,S,v,w_j}) \equiv 0\right)\right].$$

We now present the adaptations of the dictatorship gadgets for FVS and DVDP to obtain reductions from Unique Games in Sections 4.1 and 4.2, respectively. Throughout this section (as in Section 3), we fix k to be an integer, $\varepsilon, \delta > 0$ to be arbitrarily small constants and let η, d, and t be as in Theorem 4.1 (depending on k, ε and δ).

4.1 Feedback Vertex Set

We prove the following theorem from which the hardness of FVS stated in Theorem 1.1 clearly follows.

Theorem 4.2. *Assuming the Unique Games Conjecture, for any integer $k \geq 2$ and arbitrary constants $\varepsilon, \delta > 0$, given a directed graph $G(V,E)$, distinguishing between the following cases is* NP*-hard:*

- (Completeness): *there exist disjoint subsets $V_1, \ldots, V_k \subset V$ satisfying $|V_i| \geq \frac{1-2\varepsilon}{k}|V|$ and such that a subgraph induced by all but one of these subsets is acyclic.*
- (Soundness): *every induced subgraph of $8\delta|V|$ vertices contains a cycle.*

We first present the reduction in the following subsection followed by the completeness (Lemma 4.3) and soundness (Lemma 4.4) analyses.

4.1.1 Reduction

We describe a reduction from Unique Games to FVS. Let $\mathcal{L}(G(V,W,E),[R],\{\pi_{v,w}\}_{(v,w)\in E})$ be a Unique Games instance. As in Section 3.1, the FVS instance consists of two types of vertices that we refer to as *bit-vertices* and *test-vertices* and all edges are between bit- and test-vertices.

- For every $w \in W$ and $x \in [k]^R$, there is a bit-vertex $b_{w,x}$ of weight ∞.

  In other words, each $w \in W$ is replaced by a k-ary hypercube $[k]^R$ where each vertex has weight ∞ so that none of them will ever be deleted in an optimal solution.

- For every $x \in [k]^R, S \subseteq [R]$ of εR indices, $v \in V$, and $(w_1, w_2 \dots, w_{2t}) \in N(v)^{2t}$, we have a test vertex $t_{x,S,v,w_1,\dots,w_{2t}}$.

- The arcs incident to a test-vertex $t_{x,S,v,w_1,\dots,w_{2t}}$ are the following. For $j = 1,\dots,2t$,

  - there is an arc $(b_{w_j,z}, t_{v,x,S,w_1,\dots,w_{2t}})$ if $z \in C_{x,S,v,w_j}$,
  - and an arc $(t_{v,x,S,w_1,\dots,w_{2t}}, b_{w_j,z})$ if $z \in C^{\oplus}_{x,S,v,w_j} = \{z \oplus 1 : z \in C_{x,S,v,w_j}\}$.

As the bit-vertices have weight ∞, they are never deleted in an optimal solution and we can obtain an unweighted graph G (with the set T of test-vertices as vertices) with the same optimal value by having an arc between two test-vertices if there is a path between them through one bit-vertex. Theorem 4.2 therefore follows from that (i) in the proof of completeness, we can partition the test-vertices as required (Lemma 4.3) and (ii) in the proof of soundness, any solution deletes almost all test-vertices (Lemma 4.4).

4.1.2 Completeness

We prove the following.

Lemma 4.3. *If there is a labeling ρ of the Unique Games instance $\mathcal{L}$ satisfying a $(1-\zeta)$ fraction of the constraints then we can partition the test-vertices into subsets $T', T_0, \dots, T_k$ satisfying*

$$|T_j| \geq \frac{1-2\varepsilon}{k}|T|, \quad |T'| \leq 2\varepsilon|T|,$$

and for $j \in [k]$ the graph obtained by deleting $T' \cup T_j$ is acyclic.

Proof. Let ρ be such a labeling of the Unique Games instance. We now use ρ to partition the bit-vertices in k equal sized sets:

$$B_j = \{b_{w,x} : w \in W \text{ and } x_{\rho(w)} = j\} \qquad \text{for } j \in [k].$$

We say that a test-vertex $t_{x,S,v,w_1,\dots,w_{2t}}$ is good if (i) $\rho(v) \notin S$ and (ii) ρ satisfies all the edges $(v,w_1),(v,w_2),\dots,(v,w_{2t})$. Note that property (i) holds with probability at least $1-\varepsilon$ and property (ii) holds with probability at least $1-2\zeta t$. Therefore, at least a $(1-2\varepsilon)$ fraction (for ζ small enough) of the test-vertices are good. As we did for the bit-vertices, we partition the test-vertices into k equal sized sets:

$$T_j = \{t_{x,S,v,w_1,\dots,w_{2t}} : t_{x,S,v,w_1,\dots,w_{2t}} \text{ is good and } x_{\rho(v)} = j\} \qquad \text{for } j \in [k].$$

The sets are of equal size since they are partitioned according to x and whether a test-vertex is good only depends on S and $v, w_1, \ldots, w_{2t}$. Furthermore, since at least $(1-2\varepsilon)|T|$ test-vertices are good we have that

$$|T_j| \geq \frac{1-2\varepsilon}{k}|T|$$

for $j \in [k]$ and the remaining test-vertices in T' are therefore at most $2\varepsilon|T|$ many.

It remains to show that $T_j \cup T'$ defines a feedback vertex set for any $j \in [k]$. The key observation is that T_j only has incoming arcs from bit-vertices in B_j and outgoing arcs to bit-vertices in $B_{j\oplus 1}$. To see this, consider a test-vertex $t_{x,S,v,w_1,\ldots,w_{2t}}$ in T_j and let $(b_{w_i,z}, t_{x,S,v,w_1,\ldots,w_{2t}})$ be an arc. Then, by definition we have that $z \in C_{x,S,v,w_i}$. As $\pi_{v,w_i}(\rho(w_i)) = \rho(v) \notin S$,

$$z_{\rho(w_i)} = x_{\pi_{v,w_i}(\rho(w_i))} = x_{\rho(v)} = j\,,$$

and hence $b_{w_i,z} \in B_j$. The exact same argument also shows that all outgoing arcs from $t_{x,S,v,w_1,\ldots,w_{2t}}$ go to bit-vertices in $B_{j\oplus 1}$. We can therefore conclude that test-vertices in T_j have only incoming arcs from bit-vertices in B_j and outgoing arcs to bit-vertices in $B_{j\oplus 1}$.

By the key observation, we can obtain an acyclic graph by deleting all bad test-vertices and one of the sets $T_0, \ldots, T_{k-1}$ which proves the lemma.

□

4.1.3 Soundness

As we can choose the soundness parameter γ of the Unique Games Conjecture to be arbitrarily small, the following lemma says that in the proof of soundness, there is no feedback vertex set containing less than a $(1-8\delta)$ fraction of the test-vertices (or equivalently, any induced subgraph containing an 8δ fraction of the test-vertices contains a cycle).

Lemma 4.4. *If the graph has a FVS containing less than a $(1-8\delta)$ fraction of the test-vertices, then the Unique Games instance has a labeling that satisfies at least a $\delta\eta^2/(t^2k^2)$ fraction of the constraints.*

Proof. Consider a topological sort σ of the graph obtained by deleting a FVS and assume that it contains at least an 8δ fraction of the test-vertices, i. e., if we let T be the set of remaining test-vertices then

$$\Pr_{x,S,v,w_1,\ldots,w_{2t}}[t_{x,S,v,w_1,\ldots,w_{2t}} \in T] \geq 8\delta\,.$$

We shall show that it follows that there is a labeling of $\mathcal{L}$ that satisfies at least a $\delta\eta^2/(t^2k^2)$ fraction of the constraints.

With each $w \in W$, we associate the indicator function $f_w : [k]^R \to \{0,1\}$ that takes value 0 for the first half of the bit-vertices corresponding to w (according to σ) and value 1 for the remaining half. We then define the set $L[w]$ of candidate labels for every $w \in W$ as:

$$L[w] := \{i \in [R] : \mathrm{Infl}_i^d(f_w) \geq \eta\}\,.$$

By Observation 2.1, we have $|L[w]| \leq d/\eta$.

Now, for every $w \in W$, we define $\rho(w)$ to be a random label from $L[w]$ (if $L[w]$ is empty we assign any label to w) and, for every $v \in V$ we pick a random neighbor $w \in N(v)$ and define $\rho(v) = \pi_{v,w}(\rho(w))$. We shall now calculate a lower bound on the expected number of edges the labeling ρ satisfies.

We call a tuple $(v, w_1, \ldots, w_{2t})$ good if

$$\Pr_{x,S}[t_{x,S,v,w_1,\ldots,w_{2t}} \in T] \geq 4\delta\,.$$

Since T contains an 8δ fraction of the test-vertices we have that at least a 4δ fraction of the tuples are good. Consider such a good tuple $(v, w_1, \ldots, w_{2t})$ and let $T_{v,w_1\ldots,w_{2t}} = \{t_{x,S,v,\ldots,w_{2t}} \in T\}$. Suppose without loss of generality that

$$\max_{f_{w_i}(x)=0} \sigma(b_{w_i,x}) < \max_{f_{w_{i+1}}(x)=0} \sigma(b_{w_{i+1},x}) \qquad \text{for } i = 1, 2, \ldots, 2t-1\,.$$

Then if a test-vertex $t_{x,S,v,w_1,\ldots,w_{2t}}$ is in $T_{v,w_1,\ldots,w_{2t}}$ then we conclude by the arcs incident to the test-vertex that

$$\begin{cases} f_{w_{t+1}}(C_{x,S,v,w_{t+1}}) \equiv \cdots \equiv f_{w_{2t}}(C_{x,S,v,w_{2t}}) \equiv 0 & \text{if } \sigma(t_{x,S,v,w_1\ldots,w_{2t}}) \leq \max_{x: f_{w_t}(x)=0} \sigma(b_{w_t,x})\,, \\ f_{w_1}(C^{\oplus}_{x,S,v,w_1}) \equiv \cdots = f_{w_t}(C^{\oplus}_{x,S,v,w_t}) \equiv 1 & \text{otherwise, i. e., } \sigma(t_{x,S,v,w_1\ldots,w_{2t}}) > \max_{x: f_{w_t}(x)=0} \sigma(b_{w_t,x})\,. \end{cases}$$

Therefore, one of these conditions must be satisfied by at least half of the test-vertices in $T_{x,S,v,w_1,\ldots,w_{2t}}$, i. e., either

$$\Pr_{x,S}\left[\bigwedge_{j=t+1}^{2t} \left(f_{w_j}(C_{x,S,v,w_j}) \equiv 0\right)\right] \geq 2\delta$$

or

$$\Pr_{x,S}\left[\bigwedge_{j=1}^{t} \left(f_{w_j}(C^{\oplus}_{x,S,v,w_j}) \equiv 1\right)\right] = \Pr_{x,S}\left[\bigwedge_{j=1}^{t} \left((1-f_{w_j})(C_{x,S,v,w_j}) \equiv 1\right)\right] \geq 2\delta\,.$$

In either case, Theorem 4.1 gives that there exist $j \in L[w_{\ell_1}]$ and $j' \in L[w_{\ell_2}]$ for some $1 \leq \ell_1 \neq \ell_2 \leq 2t$ such that $\pi_{v,w_{\ell_1}}(j) = \pi_{v,w_{\ell_2}}(j')$.

We now follow the same argumentations as used in [1]. Overall, if we pick the tuple $(v, w_1, w_2, \ldots, w_{2t})$ at random and then w, w' at random from the set $\{w_1, \ldots, w_{2t}\}$, then with probability at least 4δ the tuple is good, with probability $1/(4t^2)$ we have $w = w_{\ell_1}$ and $w' = w_{\ell_2}$, and with probability $1/(k^2/\eta^2)$, the labeling procedure defines $j = \rho(w)$, $j' = \rho(w')$. Hence

$$\Pr_{v,w,w'}[\pi_{v,w}(\rho(w)) = \pi_{v,w'}(\rho(w'))] \geq \frac{4\delta\eta^2}{4t^2k^2}\,,$$

and (expected over the randomness of the labeling procedure)

$$\Pr_{(v,w)}[\rho(v) = \pi_{v,w}(\rho(w))] \geq \frac{\delta\eta^2}{t^2k^2}\,.$$

This shows that there exists a ρ that satisfies a $\delta\eta^2/(t^2k^2)$ fraction of the constraints. □

4.2 DAG Vertex Deletion Problem

We prove the following theorem from which the hardness of DVD stated in Theorem 1.1 follows.

Theorem 4.5. *Assuming the Unique Games Conjecture, for any integer $k \geq 2$ and arbitrary constants $\varepsilon, \delta > 0$, given a directed graph $G(V,E)$, distinguishing between the following cases is NP-hard:*

- (Completeness): *there exist disjoint subsets $V_1, \ldots, V_k \subset V$ satisfying $|V_i| \geq \frac{1-2\varepsilon}{k}|V|$ and such that a subgraph induced by all but one of these subsets has no path of length k.*
- (Soundness): *every induced subgraph of $32\delta|V|$ vertices contain a path of length $|V|^{1-\delta}$.*

We first present the reduction in the following subsection followed by the completeness (Lemma 4.6) and soundness (Lemma 4.7) analyses.

4.2.1 Reduction

We describe a reduction from Unique Games to DVD. Let $\mathcal{L}(G(V,W,E),[R],\{\pi_{v,w}\}_{(v,w)\in E})$ be a Unique Games instance. As before, it will be convenient to present the DVD instance as it consists of two types of vertices that we refer to as *bit-vertices* and *test-vertices* and all edges are between bit- and test-vertices.

- The bit-vertices are partitioned into $L+1$ bit-layers (L is selected below). Each bit-layer $\ell = 0, \ldots, L$ contains a bit-vertex $b^\ell_{w,x}$ of weight ∞ for every $w \in W$ and $x \in [k]^R$.

  In other words, each $w \in W$ is replaced by a Q-ary hypercube $[k]^R$ in each layer.
- Similarly, the test-vertices are partitioned into L test-layers. Each test-layer $\ell = 0, \ldots, L-1$ has a test-vertex $t^\ell_{x,S,v,w_1,\ldots,w_{2t}}$ of weight 1 for every $x \in [k]^R$, every subset $S \subseteq [R]$ of εR indices, every $v \in V$ and every sequence $(w_1, \ldots, w_{2t}) \in N(v)^{2t}$ of (not necessarily distinct) $2t$ neighbors of v.
- The arcs incident to a test-vertex $t^{\ell'}_{x,S,v,w_1,\ldots,w_{2t}}$ are the following. For $j = 1, \ldots, 2t$,
  - there is an arc $(b^\ell_{w_j,z}, t^{\ell'}_{x,S,v,w_1,\ldots,w_{2t}})$ if $\ell \leq \ell'$ and $z \in C_{x,S,v,w_j}$,
  - and an arc $(t^{\ell'}_{x,S,v,w_1,\ldots,w_{2t}}, b^\ell_{w_j,z})$ if $\ell > \ell'$ and $z \in C^{\oplus}_{x,S,v,w_j}$.
- Finally, L is selected so as $\delta^2 L \geq |T|^{1-\delta}$ where T is the set of test-vertices.

Similar to before, we can obtain an unweighted graph G (with the set T of test-vertices as vertices) with the same optimal value by having an arc between two test-vertices if there is a path between them through one bit-vertex. Theorem 4.5 therefore follows from that (i) in the proof of completeness, we can partition the test-vertices as required (Lemma 4.6) and (ii) in the proof of soundness, any solution deletes almost all test-vertices (Lemma 4.7) in order to avoid long paths.

4.2.2 Completeness

We show the following.

Lemma 4.6. *If there is a labeling ρ of the Unique Games instance $\mathcal{L}$ satisfying a $(1-\zeta)$ fraction of the constraints then we can partition the test-vertices into subsets $T', T_0, \ldots, T_k$ satisfying $T_j \geq \frac{1-2\varepsilon}{k}|T|, |T'| \leq 2\varepsilon|T|$, and for $j \in [k]$ the graph obtained by deleting $T' \cup T_j$ has no path of length k.*

Proof. Note that if we collapse all layers by identifying the different copies of a bit-vertex and test-vertex in different layers then the DVD instance is equivalent to the FVS instance constructed in Section 4.1. We can therefore (by the arguments of Section 4.1.2), partition the bit-vertices into k equal sized sets $B_0, B_1, \ldots, B_{k-1}$ and all but an 2ε fraction of the test-vertices into k equal sized sets $T_0, T_1, \ldots, T_{k-1}$ so that any test-vertex in T_j has only incoming arcs from bit-vertices in B_j and outgoing arcs to bit-vertices in $B_{j\oplus 1}$.

Any $j \in [k]$ therefore corresponds to a solution by removing an 2ε fraction of the test-vertices (i. e., the set T') and those test-vertices in T_j. □

4.2.3 Soundness

As we can choose the soundness parameter γ of the Unique Games Conjecture to be arbitrarily small and $\delta^2 L \geq |T|^{1-\delta}$, the proof of soundness follows from the following lemma.

Lemma 4.7. *If the Unique Games instance has no labeling that satisfies a $\delta\eta^2/(t^2k^2)$ fraction of the constraints then every induced subgraph of the bit-vertices and $32\delta|T|$ test-vertices has a path of length $\delta^2|L|$.*

Proof. As in Section 3.2.3, it will be convenient to look at the equivalent formulation of the problem where we wish to find a coloring χ that assigns a color in $\{1, 2, \ldots, k\}$ to each bit-vertex satisfying $\chi(b^{\ell}_{w,x}) \leq \chi(b^{\ell'}_{w,x})$ if $\ell \leq \ell'$ so as to minimize the number of unsatisfied test-vertices where a test-vertex $t^{\ell}_{x,S,v,w_1,\ldots,w_{2t}}$ is said to be satisfied if

$$\max_{\substack{1\leq j\leq 2t \\ z\in C_{x,S,v,w_j}}} \chi(b^{\ell}_{w_j,z}) < \min_{\substack{1\leq j\leq 2t \\ z\in C^{\oplus}_{x,S,v,w_j}}} \chi(b^{\ell+1}_{w_j,z}),$$

that is, all its predecessors are assigned lower colors than its successors.

We also generalize the concept of a lower bound on the colors assigned to most bit-vertices corresponding to $w \in W$ in each layer: define the color $\chi(w,\ell)$ of $w \in W$ and a bit-layer $\ell = 0, 1, \ldots, L$ as the maximum color that satisfies $\Pr_x[\chi(b^{\ell}_{w,x}) \geq \chi(w,\ell)] \geq 1-\delta$.

Now, with $w \in W$ and each test-layer $\ell = 0, 1, \ldots, L-1$ we associate the indicator function $f^{\ell}_w : [k]^R \to \{0,1\}$ defined as follows

$$f^{\ell}_w(x) = \begin{cases} 0 & \text{if} \quad \chi(b^{\ell+1}_{w,x}) > \chi(w,\ell), \\ 1 & \text{otherwise.} \end{cases}$$

Analogous to Claim 3.1, the key statement for the soundness analysis is the following.

Claim 4.8. *Assume the Unique Games instance $\mathcal{L}$ has no labeling satisfying a $\delta\eta^2/(t^2k^2)$ fraction of the constraints. If a 16δ fraction of the test-vertices of test-layer ℓ are satisfied, then $\chi(w,\ell) < \chi(w,\ell+1)$ for at least a 2δ fraction of the vertices in W.*

Proof. If we let T be the set of satisfied test-vertices of test-layer ℓ then (as T contains at least a 16δ fraction of the test-vertices in that layer)

$$\Pr_{x,S,v,w_1,\dots,w_{2t}}\left[\max_{\substack{1\le j\le 2t\\ z\in C_{x,S,v,w_j}}} \chi(b^{\ell}_{w_j,z}) < \min_{\substack{1\le j\le 2t\\ z\in C^{\oplus}_{x,S,v,w_j}}} \chi(b^{\ell+1}_{w_j,z})\right] \ge 16\delta\,.$$

Similar to Section 4.1.3, we call a tuple $(v,w_1,\dots,w_{2t})$ good if

$$\Pr_{x,S}\left[\max_{\substack{1\le j\le 2t\\ z\in C_{x,S,v,w_j}}} \chi(b^{\ell}_{w_j,z}) < \min_{\substack{1\le j\le 2t\\ z\in C^{\oplus}_{x,S,v,w_j}}} \chi(b^{\ell+1}_{w_j,z})\right] \ge 8\delta$$

and note that at least an 8δ fraction of the tuples are good.

By the definition of $\chi(w,\ell)$ we have $\Pr_x[\chi(b^{\ell}_{w,x}) \ge \chi(w,\ell)] \ge 1-\delta$ and therefore for a good tuple $(v,w_1,\dots,w_{2t})$,

$$7\delta \le \Pr_{x,S}\left[\max_{1\le j\le 2t} \chi(w_j,\ell) < \min_{\substack{1\le j\le 2t\\ z\in C^{\oplus}_{x,S,v,w_j}}} \chi(b^{\ell+1}_{w_j,z})\right] \le \Pr_{x,S}\left[\bigwedge_{j=1}^{2t}\left(f^{\ell}_{w_j}(C_{x,S,v,w_j}) \equiv 0\right)\right],$$

which, by Theorem 4.1, shows that either

(i) more than t of the functions are such that $\mathbb{E}[f^{\ell}_{w_j}] < \delta$ and hence $\chi(w_j,\ell+1) > \chi(w_j,\ell)$; or

(ii) there exists $1 \le \ell_1 \ne \ell_2 \le t$ and $j \in L[w_{\ell_1}]$, $j' \in L[w_{\ell_2}]$ such that $\pi_{v,w_{\ell_1}}(j) = \pi_{v,w_{\ell_2}}(j')$, where (similar to Section 4.1.3)
$$L[w] = \{i \in [R] : \mathrm{Infl}^d_i(f^{\ell}_w) \ge \eta\}\,.$$

If condition (i) holds for half the good tuples, i. e., a 4δ fraction of all tuples, then the statement follows because we can pick a vertex w in W uniformly at random by first picking a tuple $(v,w_1,\dots,w_{2t})$ at random and then picking one of the vertices $w_1,\dots,w_{2t}$ at random. With probability 4δ the tuple is good and satisfy condition (i) and (conditioned upon that fact) with probability $1/2$ the picked vertex w_j will be such that $\chi(w_j,\ell+1) > \chi(w_j,\ell)$. Therefore, we have that if condition (i) holds for half the good tuples then $\Pr_w[\chi(w,\ell+1) > \chi(w,\ell)] \ge 2\delta$ as required.

On the other hand, we shall show that the assumption of the claim (that no labeling of the Unique Games instance satisfies a $\delta\eta^2/(t^2k^2)$ fraction of the constraints) is violated if condition (ii) holds for more than half the good tuples. This follows from very similar arguments as in Section 4.1.3 (and in [1]). Indeed, for every $w \in W$, define $\rho(w)$ to be a random label from $L[w]$ and, for every $v \in V$ pick a random neighbor $w \in N(v)$ and define $\rho(v) = \pi_{v,w}(\rho(w))$. If condition (ii) holds for half the good tuples, then with probability 4δ a random tuple $(v, w_1, \ldots, w_{2t})$ is such a tuple, with probability $1/(4t^2)$ we have $w = w_{\ell_1}$ and $w' = w_{\ell_2}$ for w, w' randomly picked in the set $\{w_1, \ldots, w_{2t}\}$, and with probability $1/(k^2/\eta^2)$, the labeling procedure defines $j = \rho(w), j' = \rho(w')$. Hence (if condition (ii) holds for half the good tuples)

$$\Pr_{v,w,w'}[\pi_{v,w}(\rho(w)) = \pi_{v,w'}(\rho(w'))] \geq \frac{4\delta\eta^2}{4t^2k^2},$$

and (expected over the randomness of the labeling procedure)

$$\Pr_{(v,w)}[\rho(v) = \pi_{v,w}(\rho(w))] \geq \frac{\delta\eta^2}{t^2k^2}.$$

This shows that condition (ii) cannot hold for half the good tuples as this would imply that there is a labeling of $\mathcal{L}$ that satisfies a $\delta\eta^2/(t^2k^2)$ fraction of the constraints. □

To see how the above claim implies the lemma consider a subgraph induced by all bit-vertices and a 32δ fraction of the test-vertices and consider the smallest number of colors needed for a coloring χ to satisfy all those test-vertices.

Note that at least a 16δ fraction of the test-layers are such that at least a 16δ fraction of the test-vertices of that layer are satisfied by χ. This in turn, by the preceding claim, shows that

$$\Pr_{\ell \in [L], w \in W}[\chi(w, \ell+1) > \chi(w, \ell)] \geq 16\delta \cdot 2\delta = 32\delta^2$$

and hence there exists a $w \in W$ such that $\Pr_{\ell \in [L], w \in W}[\chi(w, \ell+1) > \chi(w, \ell)] \geq 32\delta^2$. Therefore, the coloring χ needs to use at least $32\delta^2 L$ colors to satisfy a 32δ fraction of the test-vertices or, in other words, any subgraph induced by the bit-vertices and a 32δ fraction of the test-vertices has a path of length $32\delta^2 L - 1 \geq \delta^2 L$. □

5 Discrete Time–Cost Tradeoff problem

In the Discrete Time–Cost Tradeoff problem we are given a set J of activities together with a partial order $(J, <)$. Any execution of the activities must comply with the partial order, that is, if $j < k$ activity k may not be started until j is completed. The duration of an activity depends on how much of the resources are spent on it. This tradeoff between time and cost for each job is described by a nonnegative cost function $c_j : \mathbb{R}_+ \to \mathbb{R}_+ \cup \{\infty\}$, where $c_j(x_j)$ denotes the cost to run j with duration x_j. The project duration $t(x)$ of the realization x is the makespan (length) of the schedule which starts each activity at the earliest point in time obeying the precedence constraints and durations x_j. Given a deadline $T > 0$, the deadline version of the Discrete Time–Cost Tradeoff problem (DTCT) is that of finding the cheapest realization x that obeys the deadline, i. e., $t(x) \leq T$.

Theorem 5.1. *DTCT is as hard to approximate as DVD.*

Proof. We reduce (in polynomial time) the problem of approximating DVD to that of approximating DTCT. Given an instance of DVD, i. e., an integer k and a DAG $G(V,A)$ with the vertices ordered $0,1,\dots,n-1$ according to a topological sort, consider the instance of DTCT defined as follows:

- The deadline T is set to n.
- The set J of activities contains three activities ℓ_i, m_i, r_i for each vertex $i \in V = \{0,1\dots,n-1\}$ with precedence constraints $\ell_i < c_i < r_i$ and cost functions

$$c_{\ell_i}(x) = \begin{cases} 0 & \text{if } x \geq i, \\ \infty & \text{otherwise,} \end{cases} \qquad c_{m_i}(x) = \begin{cases} 0 & \text{if } x \geq 9/10, \\ 1 & \text{otherwise,} \end{cases} \qquad c_{r_i}(x) = \begin{cases} 0 & \text{if } x > n-1-i, \\ \infty & \text{otherwise.} \end{cases}$$

  In addition, there is an activity $a_{(i,j)}$ for each arc $(i,j) \in A$ with precedence constraints $m_i < a_{(i,j)} < m_j$ and cost function

$$c_{a_{(i,j)}}(x) = \begin{cases} 0 & \text{if } x \geq j-i-\frac{9}{10}+\frac{1}{10(k-1)}, \\ \infty & \text{otherwise.} \end{cases}$$

See Figure 2 for an example of the DTCT instance corresponding to a DVD instance G with $k=3$. Note that the cost functions of ℓ_i, m_i, and r_i enforce that activity m_i has to be executed in the interval

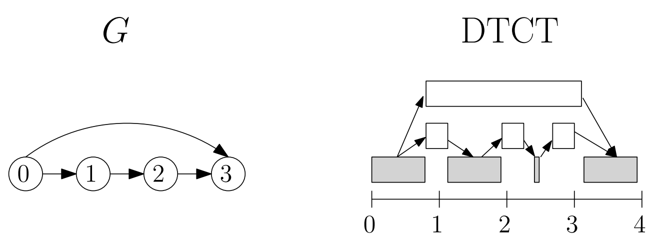

Figure 2: For each vertex $i \in V$ the activity m_i is depicted in light gray (activities ℓ_i and r_i are omitted). The activities corresponding to arcs are depicted in white. Finally, the depicted solution pays a cost of 1 for running activity m_2 in time 0.

$[i,i+1)$ and that it will require time $9/10$ unless we pay a cost of 1 which allows us to run the activity in 0 time. Furthermore, as an activity $a_{(i,j)}$ always has duration (at least)

$$j-i-\frac{9}{10}+\frac{1}{10(k-1)},$$

the start time s_j of activity m_j must be such that

$$s_j - j \geq s_i - i + \frac{1}{10(k-1)},$$

where s_i is the start time of activity i. Using the fact that an activity m_i must run in the interval $[i,i+1)$ in order to obey the deadline, it follows that we have to pay a cost of 1 for at least one activity corresponding to each path of length k. By similar arguments, it also follows that this is also sufficient for having a realization respecting the deadline. Therefore, any solution to DTCT naturally corresponds to a solution to DVD (and vice versa) by deleting those vertices corresponding to activities with a cost of 1. □

Acknowledgements

The author thanks the anonymous referees and László Babai for their valuable comments that greatly improved the presentation of the paper.

References

[1] NIKHIL BANSAL AND SUBHASH KHOT: Optimal long code test with one free bit. In *Proc. 50th FOCS*, pp. 453–462. IEEE Comp. Soc. Press, 2009. [doi:10.1109/FOCS.2009.23] 761, 763, 769, 773, 777

[2] NIKHIL BANSAL AND SUBHASH KHOT: Inapproximability of hypergraph vertex cover and applications to scheduling problems. In *Proc. 37th Internat. Colloq. on Automata, Languages and Programming (ICALP'10)*, pp. 250–261. Springer, 2010. [doi:10.1007/978-3-642-14165-2_22] 761

[3] PRABUDDHA DE, E. JAMES DUNNE, JAY B. GHOSH, AND CHARLES E. WELLS: The discrete time-cost tradeoff problem revisited. *European Journal of Operational Research*, 81(2):225–238, 1995. [doi:10.1016/0377-2217(94)00187-H] 760

[4] IRIT DINUR AND SHMUEL SAFRA: On the hardness of approximating minimum vertex cover. *Ann. of Math.*, 162(1):439–485, 2005. Preliminary version in STOC'02. See also at ECCC. [doi:10.4007/annals.2005.162.439] 761

[5] GUY EVEN, JOSEPH NAOR, BARUCH SCHIEBER, AND MADHU SUDAN: Approximating minimum feedback sets and multicuts in directed graphs. *Algorithmica*, 20(2):151–174, 1998. Preliminary version in IPCO'95. [doi:10.1007/PL00009191] 760, 761

[6] DELBERT R. FULKERSON: A network flow computation for project cost curves. *Management Science*, 7(2):167–178, 1961. [doi:10.1287/mnsc.7.2.167] 760

[7] ALEXANDER GRIGORIEV AND GERHARD J. WOEGINGER: Project scheduling with irregular costs: complexity, approximability, and algorithms. *Acta Inf.*, 41(2-3):83–97, 2004. Preliminary version in ISAAC'02. [doi:10.1007/s00236-004-0150-2] 760, 761

[8] VENKATESAN GURUSWAMI, JOHAN HÅSTAD, RAJSEKAR MANOKARAN, PRASAD RAGHAVENDRA, AND MOSES CHARIKAR: Beating the random ordering is hard: Every ordering CSP is approximation resistant. *SIAM J. Comput.*, 40(3):878–914, 2011. Preliminary versions in FOCS'08 and in CCC'09. See also at ECCC. [doi:10.1137/090756144] 761

[9] RICHARD M. KARP: Reducibility among combinatorial problems. In R. E. MILLER AND J. W. THATCHER, editors, *Complexity of Computer Computations*, pp. 85–103. Springer, 1972. [doi:10.1007/978-1-4684-2001-2_9] 760

[10] JAMES E. KELLEY, JR.: Critical-path planning and scheduling: Mathematical basis. *Operations Research*, 9(3):296–320, 1961. [doi:10.1287/opre.9.3.296] 760

[11] JAMES E. KELLEY, JR. AND MORGAN R. WALKER: Critical-path planning and scheduling. In *Papers presented at the December 1-3, 1959, eastern joint IRE-AIEE-ACM computer conference*, pp. 160–173. ACM Press, 1959. [doi:10.1145/1460299.1460318] 760

[12] SUBHASH KHOT: On the power of unique 2-prover 1-round games. In *Proc. 34th STOC*, pp. 767–775. ACM Press, 2002. [doi:10.1145/509907.510017] 761, 764

[13] FRANK THOMSON LEIGHTON AND SATISH RAO: Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms. *J. ACM*, 46(6):787–832, 1999. Preliminary version in FOCS'88. [doi:10.1145/331524.331526] 760

[14] ELCHANAN MOSSEL, RYAN O'DONNELL, AND KRZYSZTOF OLESZKIEWICZ: Noise stability of functions with low influences: Invariance and optimality. *Ann. of Math.*, 171(1):295–341, 2010. Preliminary version in FOCS'05. [doi:10.4007/annals.2010.171.295] 762, 763, 769

[15] DOOWON PAIK, SUDHAKAR M. REDDY, AND SARTAJ SAHNI: Deleting vertices to bound path length. *IEEE Trans. Computers*, 43(9):1091–1096, 1994. [doi:10.1109/12.312117] 760

[16] PRASAD RAGHAVENDRA: Optimal algorithms and inapproximability results for every CSP? In *Proc. 40th STOC*, pp. 245–254. ACM Press, 2008. [doi:10.1145/1374376.1374414] 761

[17] PAUL D. SEYMOUR: Packing directed circuits fractionally. *Combinatorica*, 15(2):281–288, 1995. [doi:10.1007/BF01200760] 760

[18] MARTIN SKUTELLA: Approximation algorithms for the discrete time-cost tradeoff problem. *Math. Oper. Res.*, 23(4):909–929, 1998. Preliminary version in SODA'97. [doi:10.1287/moor.23.4.909] 760, 762

[19] OLA SVENSSON: Hardness of vertex deletion and project scheduling. In *Proc. 15th Internat. Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX'12)*, pp. 301–312. Springer, 2012. [doi:10.1007/978-3-642-32512-0_26] 759

AUTHOR

Ola Svensson
assistant professor
École Polytechnique Fédérale de Lausanne
ola.svensson@epfl.ch
http://theory.epfl.ch/osven

ABOUT THE AUTHOR

OLA SVENSSON (not to be confused with the singer Ola Svensson) graduated from IDSIA in 2009; his advisor was Monaldo Mastrolilli. The subject of his thesis was the approximability of graph and scheduling problems. After spending two years as a postdoc with Johan Håstad at KTH Royal Institute of Technology, Sweden, he is now back in Switzerland as an assistant professor in the theory group at EPFL. Apart from doing research, he enjoys the Alps that are in complete contrast to the flat but still beautiful landscape in South Sweden where he grew up.