Better approximations of non-Hamiltonian graphs

Kazuo Iwama¹, Eiji Miyano²,*

¹ Department of Information Science, Kyoto University, Kyoto 606-01, Japan
² Department of Computer Science and Communication Engineering, Kyushu University, Fukuoka 812-81, Japan

Received 15 March 1995; received in revised form 11 June 1997; accepted 14 July 1997

Abstract

Although there have been a lot of efforts to seek nice characterization of non-Hamiltonian graphs, little progress has been made so far. An important progress was achieved by Chvátal [5, 6], who introduced the class of non-1-tough graphs (N1T) and the class of non-sub-2-factor graphs (NS2F). Both contain only non-Hamiltonian graphs and the conditions for membership can be checked in non-deterministic polynomial time. Also it is known that N1T ⊆ NS2F. Chvátal posed an open question about the complexity of those classes, i.e., whether or not they are NP-complete [6].

Recently, Bauer et al. [2] proved that N1T is NP complete. In this paper we prove: (i) NS2F is also NP-complete. (ii) NS2F − N1T is Δp-complete, namely, the former characterization for non-Hamiltonian graphs is essentially better than the latter. (iii) Those results are still true for the bounded-degree graphs.

1. Introduction

The generation of test instances to evaluate the performance of algorithms experimentally has been studied intensively from both theoretical and practical viewpoints (see, e.g., [7, 16]). For the satisfiability problem (SAT), there exists a random instance generator which can generate yes-instances (satisfiable predicates) and no-instances (unsatisfiable ones) independently [10, 11, 1], which we call a generator with known answers (GWKA). To develop a similar GWKA for the Hamiltonian circuit problem appears to be equally important. To generate yes-instances (Hamiltonian graphs) is relatively easy since the set is in NP.

On the other hand, to generate the set, NH, of non-Hamiltonian graphs efficiently appears to be very hard, since NH is co-NP-complete and it is not possible to generate a co-NP-complete set in polynomial time unless NP = co-NP. Thus, we are forced to rely upon an approximation, i.e., a certain NP subset of non-Hamiltonian graphs.

* Corresponding author. Fax: + 81 92 632 5204; e-mail: miyano@csce.kyushu-u.ac.jp.
Obviously, such a subset, \( I \), should be as large as possible because a wide range of instances can be generated. At the same time, the subset \( I \) should keep the problem's intrinsic difficulty. Suppose, for instance, that \( I \) would be a set \( I_0 \) in P. Then it would be possible to cheat the generator by developing a polynomial-time algorithm which works only for \( I_0 \cup \{y\} \), where \( \{y\} \) may be any (possibly perfect) set of yes instances.

Bauer et al. [2] proved that the set of non-\( t \)-tough graphs is NP-complete for any positive rational number \( t \). To our current purpose, this result is especially important when \( t = 1 \), since non-1-tough graphs are a rare known (non-trivial) example of NP sets which are guaranteed to include only non-Hamiltonian graphs [8]. Thus, the set of non-1-tough graphs, \( N1T \), could be a nice candidate for the approximation of non-Hamiltonian graphs in terms of GWKAs.

The main purpose of this paper is to seek a better approximation of \( N1T \) than \( N1T \). A candidate is the set called non-sub-2-factor graphs, \( NS2F \), which is known to be a proper superset of \( N1T \) and still to include only non-Hamilton graphs [6]. What we wish to do is to show that (i) \( NS2F \) is not an easy set and (ii) \( NS2F \) is essentially larger than \( N1T \).

As for (i), we prove that \( NS2F \) is NP-complete. (One should not misunderstand that since \( N1T \) is NP-hard and \( N1T \subseteq NS2F \), \( NS2F \) would be trivially NP-hard.) Since our proof works for bounded-degree graphs, the result is true for such restricted graphs and our strong conjecture is that the result is still true for constant-degree graphs. A slight modification of the proof can prove that non-\( t \)-tough graphs with bounded degrees are also NP-complete. This result strengthens [2]; in their proof being quite different from ours, vertices of unbounded degree play an important role.

As for (ii), there are no known standard methods. The fact that \( N1T \subseteq NS2F \) is clearly not sufficient because, for example, to enlarge a set with a finite number of elements is obviously possible. Actually, \( |NS2F - N1T| \) is infinite, but this fact by itself is not so important since again a set in P or NP can be enlarged infinitely with rather trivial instances [13]. In this paper, we prove that \( NS2F - N1T \) is \( D^P \)-complete. \( D^P \) is the class of languages which can be expressed by the difference \( I_1 - I_2 \) of two NP-sets \( I_1 \) and \( I_2 \). So, \( D^P \)-completeness of \( I_1 - I_2 \) assures the following: Even if there would be another (possibly better than \( I_1 \)) NP set \( I'_1 \), \( I'_1 - I_2 \) can be reduced to \( I_1 - I_2 \) easily, which means \( I_1 - I_2 \) includes all "important" members. That is why we claimed that \( NS2F \) is a significantly better approximation than \( N1T \). This seems to be a nice application of \( D^P \)-completeness.

2. \( t \)-tough, sub-2-factor and Hamiltonian graphs

Let \( G \) be a simple undirected graph. \( V \) and \( E \) denote the sets of vertices and edges, respectively. \( w(G) \) denotes the number of connected components of \( G \). \( NH \) denotes the set of all non-Hamiltonian graphs. Let \( t \) be a positive rational number. \( N1T \) and \( NS2F \) denote the sets of all non-\( t \)-tough and all non-sub-2-factor graphs, respectively. The following definitions and propositions are due to [5, 6, 8].
Definition 1. A graph $G$ is said to be $t$-tough if

$$ t \cdot w(G - S) \leq |S| $$

for any subset $S$ of $V$ such that $w(G - S) > 1$. (Here, $G - S$ is a subgraph of $G$ that is induced by the vertex set $V - S$. Similar notations will appear also below. In the original definition [5], $t$ may be real. But to discuss the NP-hardness, it needs to be rational [2].)

Definition 2. A graph $G$ is said to be sub-2-factor if

$$ w(T) \leq |S| + \sum_{Q_i} \left\lfloor \frac{\text{edge}(Q_i, T)}{2} \right\rfloor $$

for any partition of $V$ into disjoint subsets $R, S$ and $T$ (i.e., $T = G - R - S$) such that $T \neq V$, where $Q_i$ denotes each subset of vertices connected in $R$. edge$(Q_i, T)$ denotes the number of edges which have one endpoint in $Q_i$ and the other in $T$.

Example 1. See Fig. 1. This graph is (a) 1-tough but (b) not 2-tough. The reason: (a) One can easily check that there is no way of creating more connected components than the number of the removed vertices (e.g., if $v_1$ and $v_2$ are removed, then the rest of vertices, $v_3, v_4$, and $v_5$, constitute a single connected component). (b) By removing $v_2, v_3$, and $v_4$ ($= S$), the rest of the graph breaks down to two connected components, namely, $2 \cdot w(G - S) > |S|$ is satisfied.

Example 2. The graph shown in Fig. 2 is not sub-2-factor. Consider the partition of the vertices into $R, S$, and $T$ as shown in the figure. ($Q_1$ and $Q_2$ denote subsets of connected vertices in $R$.) Now it turns out that: (i) $|S| = 1$. (ii) $w(T)$ is equal to the number of the vertices labeled $T$, namely, four. (iii) The number of edges which connect $Q_1$ and $T$ is three. That is the same for $Q_2$ and $T$. Therefore,

$$ \sum_{Q_i} \left\lfloor \frac{\text{edge}(Q_i, T)}{2} \right\rfloor = \left\lfloor \frac{3}{2} \right\rfloor + \left\lfloor \frac{3}{2} \right\rfloor = 2. $$
Thus,
\[ w(T) > |S| + \sum_{T} \left( \frac{\text{edge}(Q_i, T)}{2} \right). \]

One can also check that this graph is 1-tough.

**Proposition 1** (Chvátal [5]). \( N1T \not\subset NH \).

Note that \( N1T \) is a special case of \( NS2F \), i.e., the case of \( R = \emptyset \). It is also known that:

**Proposition 2** (Chvátal [6]). \( N1T \not\subset NS2F \not\subset NH \).

**Remark 1.** It immediately follows from the definition that if \( t \leq t' \) then \( NtT \subset Nt'T \). Therefore, \( NtT \) for some \( t > 1 \) would be a better approximation of \( NH \) if it is in \( NH \). Unfortunately, that is not the case: Consider a simple circuit graph. Then, by removing non-adjacent two vertices as \( S \), we can separate two connected components. Hence, \( (1 + \varepsilon) \cdot w(G - S) \ll |S| \) for any positive number \( \varepsilon \), i.e., the simple circuit graph is not \((1 + \varepsilon)\)-tough, but is obviously a Hamiltonian graph.

**Remark 2.** Enomoto et al. [8] conjecture that every 2-tough graph is Hamiltonian. Fig. 3 illustrates the relation of those classes. (\( NH \) is the outside of the circle denoted by “Hamiltonian”. \( N1T \) is also the outside of the circle denoted by “1-tough” and so on.)
The class $D^p$, first introduced by [14], consists of all the languages that can be expressed in the form $L_1 - L_2$, where both $L_1$ and $L_2$ are in NP, or equivalently, in the form $L_1 \cap L_2$, where $L_1$ is in NP and $L_2$ in co-NP. It turns out that $D^p$ includes NP $\cup$ co-NP and is included by $D^p_2$. Both inclusions are proper unless NP = co-NP. There are three types of problems [4, 15] that seem to be in $D^p$ - (NP $\cup$ co-NP). The first type is called the exact answer problem, e.g., the exact $K$-clique problem. This problem asks, given a graph $G$ and an integer $K$, whether the maximum clique size is exactly $K$. The second type is called the criticality or maximization problem, e.g., the critical satisfiability problem: Given a CNF formula $f$, the problem asks if it is the case where the formula $f$ itself is unsatisfiable but it becomes satisfiable if we remove any single clause form $f$. The third one is the uniqueness problem, e.g., the unique satisfiability problem which asks if the number of satisfying truth assignments is exactly one [3]. The first two problems are known to be $D^p$-complete [14, 15], but as for the third problem, it is open.

3. Main theorems

**Theorem 1.** The problem of determining whether a given graph is in $NS2F$ is NP-complete.

See Section 4 for the proof. 3-SAT is reduced into this problem. It will turn out that the reduction transforms yes-instances of 3-SAT into $N1T$ graphs and no-instances into $NS2F$ (i.e., sub-2-factor) graphs. Thus, it also proves that $N1T$ is NP-complete.

**Theorem 2.** The problem of determining whether a given graph is in $NS2F - N1T$ is $D^p$-complete.

**Theorem 3.** Theorems 1 and 2 also hold for graphs who degree $\leq 15$.

**Theorem 4.** Let $t = a/b$, where both $a$ and $b$ are integers such that $a, b \geq 1$. Then the problem of determining whether a given graph is in $N1T$ is NP-complete for graphs whose degree $\leq 6a + 9b - 3$.

Recall that a graph $G$ is in $NS2F$ if there are disjoint subsets $R$ and $S$ of $V$ (let $T = V - R - S$) such that

$$w(G - R - S) > |S| + \sum_{Q \in V} \left[ \frac{\text{edge}(Q, T)}{2} \right].$$

Let $NS2F(K)$ be the set of (non-sub-2-factor) graphs $G$ for which we can satisfy this condition for some $S$ and $R$ such that $|R| \leq K$. Namely, $NS2F(0)$ is equal to $N1T$. Theorem 2 says that $\bigcup_{K \geq 1} NS2F(K)$ is $D^p$-complete. We have the following
conjecture that claims \( \bigcup_{K \geq 1} NS_2F(K) \) constitutes an infinite hierarchy with respect to \( K \):

**Conjecture.** \( NS_2F(K + 1) - NS_2F(K) \) is \( \text{D}^p \)-complete for any \( K \) such as \( K = 1, 2, \ldots, \) constant, \( \log n, \ldots, cn \) (for some constant \( c < 1 \)).

In [12], we study the complexity of *Resolution*, which is a proof system for the set of unsatisfiable CNF predicates, and show that it constitutes a similar hierarchy as above. Let *Res*(\( K \)) be the set of predicates that are proved by Resolution with at most \( K \) repetitions of clauses. Iwama and Miyano prove that *Res*(\( K \)) - *Res*(\( K - 1 \)) is \( \text{D}^p \)-complete for any constant \( K \geq 1 \). The result is similar to the \( \text{D}^p \)-completeness of the exact \( K \)-clique problem. However, note that *Res*(\( K \)) - *Res*(\( K - 1 \)) is \( \text{D}^p \)-complete (and also the conjecture above) even if \( K = 1 \), while the exact \( K \)-clique problem is obviously in \( \text{P} \) if \( K \) is constant.

4. **Proof of Theorem 1**

We show that the problem of deciding, given graph \( G = (V, E) \), whether or not there are partitions of vertices into \( R, S \) and \( T \) such that \( w(G - R - S) > |S| + \sum_Q \text{edge}(Q, T)/2 \) is NP-complete for graphs whose degree \( \leq 15 \). Since we can check the condition in polynomial time for particular \( R, S \) and \( T \), the problem is in \( \text{NP} \). To show its completeness, 3-SAT is reduced to this problem. Given a predicate \( f \), we construct the graph \( G \) satisfying the following conditions (i) and (ii): (i) There exists a subset \( S \subseteq V \) such that \( w(G - S) > |S| \) if \( f \) is satisfiable. (ii) If \( f \) is not satisfiable, the condition \( w(G - R - S) > |S| + \sum_Q \text{edge}(Q, T)/2 \) cannot be satisfied no matter how we select \( R, S \) and \( T \). The NP-completeness proof for the vertex cover problem [9] has hinted to the following proof.

**Remark.** The above reduction also claims the NP-completeness of *N1T*.

Suppose that the predicate \( f \) uses \( n \) variables, \( U = \{u_1, u_2, \ldots, u_n\} \), and contains \( m \) clauses, \( C = \{c_1, c_2, \ldots, c_m\} \), where the \( j \)th clause includes three laterals \( x_{j,1}, x_{j,2} \) and \( x_{j,3} \). The graph \( G \) consists of three subgraphs \( SG_1 \), \( SG_2 \) and \( SG_3 \). The first subgraph \( SG_1 \) is associated with the variable set \( U \), and \( SG_2 \) with the clause set \( C \). \( SG_3 \) plays an important role to manage the condition for *N1T* and *NS2F*.

\( SG_1 \) is further divided into \( n \) (= the number of variables in \( f \) ) components \( G_u_i = (V_{u_i}, E_{u_i}) \) for \( i = 1, \ldots, n \), corresponding to variable \( u_i \) of \( f \). As illustrated in Fig. 4, \( G_{u_1}, \ldots, G_{u_n} \) are in the same form and each \( G_u \) has six vertices, \( u_{i,0}, u_{i,1}, u_{i,2}, \bar{u}_{i,0}, \bar{u}_{i,1}, \bar{u}_{i,2} \). Edges exist between every two of the three vertices \( u_{i,0}, u_{i,1} \) and \( u_{i,2} \), and similarly, every two of the three vertices \( \bar{u}_{i,0}, \bar{u}_{i,1}, \bar{u}_{i,2} \). Another edge connects \( u_{i,0} \) and \( \bar{u}_{i,0} \).

The second subgraph \( SG_2 \) has \( m \) (= the number of clauses) components. Fig. 5 illustrates one of them. For each clause \( c_j \) \( (j = 1, \ldots, m) \) of \( f \), we introduce a
component, $G_{c_j} = (V_{c_j}, E_{c_j})$. $G_{c_j}$ has nine vertices, $x_{j,1,0}, x_{j,1,1}, x_{j,1,2}, x_{j,2,0}, x_{j,2,1}, x_{j,2,2}, x_{j,3,0}, x_{j,3,1}, x_{j,3,2}$. Edges are drawn between every two of the three vertices $x_{j,1,0}, x_{j,2,0}$ and $x_{j,3,0}$, and also every two of the three vertices $x_{j,k,0}, x_{j,k,1}$ and $x_{j,k,2}$ for each $k$.

The third subgraph $SG_3$ is illustrated in Fig. 6. $SG_3$ consists of (i) $n + m$ complete subgraphs, denoted by $A_1, A_2, \ldots, A_{n+m}$, of three vertices, (ii) $n + m - 1$ subgraphs denoted by $B_1, B_2, \ldots, B_{n+m-1}$, each of which has two vertices but no edges between them, and (iii) a special subset $B_0$ which has three vertices without any edge among them. We introduce the complete bipartite connection between $A_1$ and $B_0$, namely, an edge exists between any vertices $v$ and $u, v$ and $A_1$ and $u$ in $B_0$. Similar bipartite connections exist between $B_0$ and $A_2$, between $A_2$ and $B_1$, and so on.
There are also the two following sets of edges among those three subgraphs. The first set of such edges connects \(SG_1\) with \(SG_2\). For simplicity of description, we consider the following example as \(f\):
\[
f = \cdots (u_3 + \overline{u_5} + \overline{u_8}) \cdots,
\]
where \((u_3 + \overline{u_5} + \overline{u_8})\) is the \(j\)th clause. The three vertices \(x_{j,1,0}\), \(x_{j,2,0}\) and \(x_{j,3,0}\) are associated with \(u_3\), \(\overline{u_5}\) and \(\overline{u_8}\), respectively. Namely, we connect between \(u_{3,0} \in V_{u_3}\) of \(SG_1\) and \(x_{j,1,0} \in V_{c_j}\) of \(SG_2\). Similarly we connect \(\overline{u_5,0}\) with \(x_{j,2,0}\) and \(\overline{u_8,0}\) with \(x_{j,3,0}\) (see Fig. 7).

The second set of such edges connects between \(SG_3\) and \(SG_1 \cup SG_2\). The complete bipartite connection is provided between the three vertices of \(A_i\) and six vertices of \(G_{u_i}\) for each \(i = 1, \ldots, n\). Also the complete bipartite connection exists between \(A_{n+i}\) and \(V_{c_j}\) for each \(j = 1, \ldots, m\).

The whole structure of \(G\) is illustrated in Fig. 8.

**Lemma 1.** If \(f\) is satisfiable, then the graph \(G\) is in \(N \setminus T\) (i.e., in \(NS2F\)).

**Proof.** We shall show that if \(f\) is satisfiable then we can find a subset \(S\) of \(V\) such that \(w(G - S) > |S|\). As \(S\), we take the following vertices: (i) For each \(i\), one of \(u_{i,0}\)
and $\overline{u_i,0}$ of SG$_1$ is taken ($n$ vertices in total) depending on the truth assignment that makes $f$ true. If $u_i = true$ then $\overline{u_i,0}$ is taken; otherwise, $\overline{u_i,0}$ is taken. (ii) As for SG$_2$, two vertices of $x_{j,1,0}$, $x_{j,2,0}$ and $x_{j,3,0}$ are taken ($2m$ vertices in total). Which one of $x_{j,1,0}$, $x_{j,2,0}$ and $x_{j,3,0}$ is not taken is determined again by the truth assignment. Namely, the one corresponding to the literal which is true under the truth assignment remains. We can find at least one such vertex since every clause is true; choose one arbitrary vertex if two or more exist. (iii) As for SG$_3$, for $i = 1, \ldots, n + m$, all the three vertices of $A_i$ are taken ($3(n + m)$ in total).

Let us calculate how many connected components (cc’s in short from now on) $G$ is decomposed into: (1) In SG$_3$, all the $A_i$’s are taken (put into $S$). Since every vertex $x$ in $B_i$’s becomes separated, there are $2(n + m) + 1$ cc’s. (2) Let us consider SG$_1$. Suppose that $u_{i,0}$ was taken in step (i). Then we can obtain two cc’s, $\{u_{i,1}, u_{i,2}\}$ and $\{\overline{u_{i,0}}, \overline{u_{i,1}}, \overline{u_{i,2}}\}$ by the following reason: Recall that taking $u_{i,0}$ means that literal $u_i$ is false by the truth assignment. So, all the vertices in SG$_2$ connected to $\overline{u_{i,0}}$ were taken in step (ii) above. Thus, we can create $2n$ new cc’s from SG$_1$. (3) Finally consider SG$_2$. In step (ii) above, if for example $x_{j,1,0}$ remains (is not taken) then the above rule guarantees that the vertices in SG$_1$ once connected with $x_{j,1,0}$ have been taken. Henceforth, after removing the vertices $x_{j,2,0}$ and $x_{j,3,0}$ we can obtain three independent cc’s, $\{x_{j,1,0}, x_{j,1,1}, x_{j,1,2}\}$, $\{x_{j,2,1}, x_{j,2,2}\}$ and $\{x_{j,3,1}, x_{j,3,2}\}$. Thus, we have created $3m$ cc’s in total.

As a result, one can see that

$$|S| = 3(n + m) + n + 2m = 4n + 5m,$$

$$w(G - S) = 2(n + m) + 1 + 2n + 3m = 4n + 5m + 1,$$

which satisfies the condition of N1T.

It remains to prove that, if $f$ is not satisfiable then $w(T) > |S| + \sum_{Q} \lfloor$edge$(Q, T)$/$2\rfloor$ cannot be satisfied for any vertex partition into $R$, $S$ and $T$. First of all, we divide the reduced graph $G$ into $n + m$ subgraphs, denoted by $\mathcal{G}_1, \ldots, \mathcal{G}_n$, $\mathcal{G}_{n+1}, \ldots, \mathcal{G}_{n+m}$, as follows: $\mathcal{G}_1$ is induced by $A_1$ and the special vertex set $B_0$ of SG$_3$ and $G_{u_i}$ of SG$_1$ (see Fig. 9). All the edges among those vertices remain. Each $\mathcal{G}_i$ for $i = 2, \ldots, n$ is induced by $A_i, B_{i-1}$ and $G_{u_i} \subseteq$ SG$_1$ (Fig. 10 illustrates $\mathcal{G}_2$). Each $\mathcal{G}_{n+i}$ for $j = 1, \ldots, m$ is induced by $A_{n+j}, B_{n+j-1}$ and $G_{x_{j,0}} \subseteq$ SG$_2$ (Fig. 11 illustrates $\mathcal{G}_{n+1}$). The set of vertices in $\mathcal{G}_j$ is denoted by $\mathcal{V}_j$. Let $\mathcal{G}_j = \mathcal{V}_j \cap R$, $\mathcal{S}_j = \mathcal{V}_j \cap S$, $\mathcal{T}_j = \mathcal{V}_j \cap T$ and $\mathcal{E}_{ij} = \mathcal{V}_j \cap \mathcal{Q}_i$. $\mathcal{G}_i \cap \mathcal{G}_j$ denotes the subgraph induced by all the vertices in $\mathcal{G}_i$ and $\mathcal{G}_j$ which also include the edges between $\mathcal{G}_i$ and $\mathcal{G}_j$ (if any).

Let $\pi$ be a partition of $V$ into $R$, $S$ and $T$. Then we define $cost_\pi(G) = |S| + \sum_{Q} \lfloor$edge$(Q, T)$/$2\rfloor$ and call it the cost of $G$ for $\pi$. If $\pi$ is clear, we simply write $cost(G)$ and say the cost of $G$. Similarly, for each subgraph $\mathcal{G}_j$, we define the cost of $\mathcal{G}_j$ for $\pi$; namely, $cost_\pi(\mathcal{G}_j) = |\mathcal{S}_j| + \sum_{\mathcal{E}_{ij}} \lfloor$edge$(\mathcal{E}_{ij}, \mathcal{T}_j)$/$2\rfloor$. It should be noted that although $T$ must not be equal to $V$, $T$ may include all the vertices of some $\mathcal{G}_j$. The following lemma is easy but important:
Fig. 9. $\mathcal{G}_1$.

Fig. 10. $\mathcal{G}_2$.

Fig. 11. $\mathcal{G}_{n+1}$.
Lemma 2. For any vertex partition \( \pi \), the following two inequalities hold:

\[
\begin{align*}
\sum_{j=1}^{n+m} w_{\pi}(\mathcal{T}_j) & \leq w_{\pi}(T), \\
\sum_{j=1}^{n+m} \text{cost}_{\pi}(\mathcal{G}_j) & \geq \text{cost}_{\pi}(G),
\end{align*}
\]

where \( w_{\pi}(T) \) and \( w_{\pi}(\mathcal{T}_j) \) denote the number of connected components of \( G \) induced by \( T \)-vertices and the number of connected components of \( \mathcal{T}_j \) induced by \( \mathcal{T}_j \) vertices, respectively.

Proof. Let \( G' \) be the graph which consists of \( \mathcal{G}_1, \ldots, \mathcal{G}_{n+m} \), namely, \( G' \) is the same as \( G \) but edges among those subgraphs are missing. Then one can see that the right-hand side of (3) is equal to the number of cc's of \( G' \) for the same \( T \). It is obvious that removing edges does not decrease the number of cc's and, therefore, \( w_{\pi}(T) \) of \( G' \) is greater than or equal to \( w_{\pi}(T) \) of \( G \). Thus (3) holds. Also, the right-hand side of (4) is equal to the cost of \( G' \). Again the cost does not increase (i.e., the number of edges between \( T \) and \( R \) can only decrease) by removing edges, so (4) holds as well.

Note that we are now trying to show that there is no vertex partition which creates more cc's than its cost. In the rest of the proof, we shall rule out several partitions being effective in this sense for subgraphs. Then, for the whole graph, we obtain the most effective partition, which is essentially the same as the partition we used in Lemma 1, under the condition that each subgraph \( \mathcal{G}_i \) has at least one vertex not in \( T \). Recall that the key point in the proof of Lemma 1 is that at least one vertex out of every two vertices connected by edges between \( SG_1 \) and \( SG_2 \) is in \( S \). It should be noticed that such a partition was possible only because \( f \) is satisfiable. This time, \( f \) is unsatisfiable and that most effective partition still cannot produce enough number of cc's. Finally, we discuss the case that all the vertices of some \( \mathcal{G}_i \) are in \( T \).

Now let us take a look at the three graphs in Figs. 9–11. One can see that \( \mathcal{G}_2 \) is minimum among the three and the other graphs \( \mathcal{G}_1 \) and \( \mathcal{G}_{n+1} \) can be obtained by adding some vertices (and edges) to \( \mathcal{G}_2 \). So, we shall have a detailed discussion on how to decompose vertices for the simplest \( \mathcal{G}_2 \) first and then the discussion for the other graphs will be simplified since we only have to observe how the excessive vertices are further decomposed.

Lemma 3. Let \( \pi \) be a vertex partition such that \( \mathcal{G}_2 \) has at least one vertex not in \( T \). Then, \( w_{\pi}(\mathcal{G}_2) \leq \text{cost}_{\pi}(\mathcal{G}_2) \). Same for \( \mathcal{G}_i \) for \( 3 \leq i \leq n \).

Proof. The following five cases on the partition of the three vertices of \( A_2 \) in \( SG_3 \) of \( \mathcal{G}_2 \) will be considered (see Fig. 10 again):

Case 1: There is at least one \( T \)-vertex in \( A_2 \). Since that \( T \)-vertex is connected to all the other vertices of \( \mathcal{G}_2 \), \( w(\mathcal{G}_2) \) is one. One can see that the cost is also at least
one as follows: (1) If \( \mathcal{G}_2 \) has at least one \( S \)-vertex, then \( \text{cost}(\mathcal{G}_2) \) is at least one. (2) No vertices are in \( S \) and one or two vertices in \( A_2 \) are in \( T \) (so two or one vertex is in \( R \)). Then, \( \text{cost}(\mathcal{G}_2) \geq 1 \) because there are at least two \( T-R \)-edges. (3) Suppose that all the three vertices of \( A_2 \) are in \( T \). Then, from the condition, there must be at least one vertex of \( B_1 \cup G_{u_2} \) is in \( R \). Since this \( R \)-vertex has at least three \( T-R \)-edges, \( \text{cost}(\mathcal{G}_2) \) is at least one. Thus, \( w(\mathcal{F}_2) \leq \text{cost}(\mathcal{G}_2) \).

Case 2: All of the three vertices of \( A_2 \) are in \( S \) (cost three). Now, the two vertices of \( B_1 \) are isolated (two \( cc \)'s). So putting the two vertices of \( B_1 \) into \( T \) is the worst case for the condition \( w(\mathcal{F}_2) \leq \text{cost}(\mathcal{G}_2) \) since otherwise, for example, if one of the two \( B_1 \)-vertices is in \( R \) then \( w(\mathcal{F}_2) \) decreases and \( \text{cost}(\mathcal{G}_2) \) does not (if that \( B_1 \)-vertex is in \( S \) then the cost even increases). Also \( G_{u_2} \) shown in Fig. 10 is separated. Then, we can make the following claim:

**Claim 1.** At most, \( h \) \( cc \)'s can be separated from \( G_{u_2} \) at the cost of \( h-1 \) for any vertex partition.

Thus \( w(\mathcal{F}_2) \leq \text{cost}(\mathcal{G}_2) \) follows if we take \( A_2 \) and \( B_1 \) into account.

**Proof of the Claim 1.** The following eight cases are considered. One can see easily that we can obtain at most two \( cc \)'s from \( G_{u_2} \) by whatever partition \( \pi \).

1. All of the six vertices of \( G_{u_2} \) are in \( T \). Then, we get one \( cc \) \( G_{u_2} \) (cost 0).
2. One of them is in \( S \) and the other five vertices in \( T \). If one of the inner two vertices \( u_{2,0} \) and \( \overline{u}_{2,0} \) is in \( S \), then \( G_{u_2} \) is divided into two \( cc \)'s and the cost is one. Otherwise, taking one of the outer four vertices \( u_{2,1}, \overline{u}_{2,1}, u_{2,2}, \overline{u}_{2,2} \) into \( S \) provides no new \( cc \)'s.
3. Two or more of them are in \( S \) and the others in \( T \). Then, the cost is two or more. On the other hand, the number of \( cc \)'s is at most two.
4. One of them is in \( R \) and the others in \( T \). (i) If one of \( u_{2,0} \) and \( \overline{u}_{2,0} \) is in \( R \), then, as (2)-(i), we obtain two \( cc \)'s at the cost of one (\( \frac{3}{2} \)). (ii) If one of \( u_{2,1}, u_{2,2} \), \( \overline{u}_{2,1}, \overline{u}_{2,2} \) is in \( R \), then the cost is \( \frac{3}{2} \) but we can separate no more \( cc \)'s.
5. Two of them are in \( R \) and the others in \( T \). (i) If the two inner vertices, \( u_{2,0} \) and \( \overline{u}_{2,0} \), are in \( R \), then the cost and the number of \( cc \)'s are both two. (ii) If, for example, the inner vertex \( u_{2,0} \) and the outer \( u_{2,1} \) are in \( R \), then the cost is one but the number of \( cc \)'s are two. (iii) If the outer two vertices are in \( R \), then the cost and the number of \( cc \)'s are both one.
6. Three of them are in \( R \) and the others in \( T \). If \( u_{2,0}, u_{2,1}, u_{2,2} \) (or \( \overline{u}_{2,0}, \overline{u}_{2,1}, \overline{u}_{2,2} \)) are in \( R \), then the \( cc \) is one at no cost. However, for any other case, the cost is at least two.
7. Four or more of them are in \( R \) and the others in \( T \). Considering that the degree of \( T \)-vertices is two, we cannot obtain more \( cc \)'s than its cost.
8. We are finished the cases such that the six vertices are all in \( T \) (Case (1)), all in \( S \) (Case (2)), all in \( R \) (Case (7)), in \( S \) or \( T \) ((2) and (3)) and in \( R \) or \( T \) ((4) through (7)). So remaining cases are: (i) the vertices are in \( S \) or \( R \) and (ii) in \( S \) or \( R \).
or $T$. Note that both cases can be considered as either Case (2) or (3) in which some $T$ vertices are changed into $R$-vertices. Thus as Case (3), we can obtain no more than two cc’s but the cost is at least one ($|S| \geq 1$).

**Case 3:** One of the three vertices of $A_1$ is in $R$ and the other two in $S$. Similarly, $B_1$ is isolated. However, this case is a little different from Case 2: If $B_1$ has $T$-vertices, then there are $T$-$R$-edges between $A_2$ and $B_1$. Let us consider the subgraph $A_2 \cup B_1$: (i) If the two vertices of $B_1$ are in $T$, then the cost of $A_2 \cup B_1$ is three and there are two cc’s (two vertices of $B_1$). (ii) If one of them is in $T$ and the other in $R$, then the cost is two and only one cc is separated. (iii) If one of them is in $T$ and the other in $S$, then the cost is three. (iv) If both of them are in $R$, or (v) Both in $S$, then the cost is two or more but there is no cc. Thus, in any case $w(F_2) \leq \min(3, 1)$ by Claim 1.

**Case 4:** Two of them $A_1$ are in $R$ and one in $S$. If $z$ (at least one by the condition of the lemma) vertices of $B_1 \cup G_{u_1}$ are in $T$, then $\min(3, 1) + 1 = \min(4, 2) + 1$ since there are at least $2z$ $T$-$R$ edges but $w(F_2) \leq z$.

**Case 5:** All of them of $A_1$ are in $R$. If $z > 1$ vertices of $B_1 \cup G_{u_1}$ are in $T$, then $\min(3, 1) + 1 = \min(4, 2) + 1$ since each of those vertices has at least three edges connecting $A_2$-vertices now in $R$ but $w(F_2) \leq z$.

We next consider $\mathcal{G}_{n+1}$ (and $\mathcal{G}_{n+2}, \ldots, \mathcal{G}_{n+m}$ are the same) that differs from $\mathcal{G}_2, \ldots, \mathcal{G}_n$ in $G_{c_1}$ (see Fig. 11).

**Lemma 4.** Let $\pi$ be a vertex partition such that $\mathcal{G}_{n+1}$ has at least one vertex not in $T$. Then, $w_{\pi}(F_{n+1}) \leq \min(4, 2) + 1$. Same for $\mathcal{G}_i$ for $n + 2 \leq i \leq n + m$.

**Proof.** Recall that $\mathcal{G}_{n+1}$ is the same as $\mathcal{G}_2$ but $G_{c_1}$ in $G_{n+1}$ has further one complete graph with three vertices, e.g., $x_{1,1,0}$, $x_{1,1,1}$ and $x_{1,1,2}$. We will focus our attention to the vertex partition of that complete graph, say, $H$. Note that at most one cc can be separated from any complete graph by whatever partition $\pi$, which means the number of cc’s can increase at most one compared to $\mathcal{G}_2$ in Lemma 3. If $H$ has at least one vertex in $R$ or $S$, then the cost of $H$ is also at least one, i.e., the lemma follows from Lemmas 2 and 3. Hence, suppose that all of the vertices of $H$ are in $T$. If $H$ is connected to any other $T$-vertices (no cc-increase), or to two or more $R$-vertices (one cost-increase), then the lemma holds. Thus, $A_{n+1} \subseteq S$ and $B_n \subseteq T$ in Case 2 of Lemma 3 is the only remaining case to be observed. Now there are two possibilities: (i) $x_{1,2,0}$ and $x_{1,3,0}$ are both in $S$. (ii) One of them is in $S$ and the other in $R$. Both cases are similar to before; the number of cc’s from $G_{u_1}$ is at most the same as its cost, or $w_{\pi}(F_2) \leq 4$ and $\min(3, 1) + 1 = \min(4, 2) + 1$. Hence, $w_{\pi}(F_{n+1}) \leq 5$ and $\min(3, 1) + 1 = \min(4, 2) + 1$.

We next consider $\mathcal{G}_1$ in which $B_0$ has three vertices, one more vertex than $B_1$.

**Lemma 5.** Let $\pi$ be a vertex partition such that $\mathcal{G}_1$ (Fig. 9) has at least one vertex not in $T$. Then, $w_{\pi}(F_2) \leq \min(4, 2) + 1$. If (i) $A_1$ has at least one vertex not in $S$, or (ii) $B_0$ has at least one vertex not in $T$.
Proof. Again the number of cc's can increase at most one compared to Lemma 3. So, if at least one vertex, say, \( v_0 \) in \( B_0 \) is in \( R \) or \( S \), then the lemma holds. The reason is that we can consider that \( v_0 \) is added to \( B_1 \) of Lemma 3, which increases the cost by one compared to Lemma 3. Thus, we have nothing to prove for the case (ii).

Now let us discuss the case (i) under the condition that all three vertices in \( B_0 \) is in \( T \). If \( A_1 \) includes \( T \)-vertices then the number of cc's does not increase from Lemma 3. If \( A_1 \) includes two or more \( R \)-vertices, then the cost increases. Namely, both cases are immediate from Lemma 3. So, the only remaining case is that one of the three vertices of \( A_1 \) is in \( R \) and the other two in \( S \): If \( G_{u_i} \) has no \( T \)-vertices, then \( T \)-vertices only exist in \( B_0 \) and \( A_1 \). Otherwise, i.e., if there are some \( T \)-vertices in \( G_{u_i} \), then its cost is at least \( h-1 \) by Claim 1 of Lemma 3. Now, there is at least one \( T-R \)-edge between \( G_{u_i} \) and \( A_1 \) in addition to the three ones between \( B_0 \) and \( A_1 \). Hence, \( \text{cost}(\mathcal{G}) \geq (h-1) + \left\lceil \frac{h}{2} \right\rceil + 2 = h + 3 \) and \( w(\mathcal{G}) \leq h + 3 \). □

By Lemmas 2–5, we can conclude that \( w(T) \leq \text{cost}(G) \) if at least one vertex in \( A_1 \) is not in \( S \) or at least one vertex in \( B_0 \) is not in \( T \). Hence, we shall now consider the case that all \( A_1 \)-vertices are in \( S \) and all \( B_0 \)-vertices are in \( T \).

Lemma 6. Suppose that \( \pi \) is a vertex partition such that each \( \mathcal{G}_i \) has at least one vertex not in \( T \), \( A_1 \subseteq S \) and \( B_0 \subseteq T \). Then \( w_\pi(T) \leq \text{cost}_\pi(G) \) nevertheless hold if \( A_i \notin S \) for some \( i \neq 1 \) or \( B_j \notin T \) for some \( j \neq 0 \).

Proof. As before, let \( G' \) be the graph which consists of \( \mathcal{G}_1, \ldots, \mathcal{G}_{n+m} \). We consider \( \mathcal{G}_i \)'s adjacent subgraph \( \mathcal{G}_2 \). (i) If \( A_2 \) has at least one \( R \)-vertex (i.e., \( A_2 \notin S \)), then, for the same vertex partition, \( \text{cost}(G) = \text{cost}(G') + 1 \) since \( G \) has three \( T-R \)-edges between \( \mathcal{G}_1 \) and \( \mathcal{G}_2 \) but \( G' \) has no edge between them. Recall that \( w_\pi(\mathcal{G}) \leq \text{cost}_\pi(\mathcal{G}) + 1 \) for any \( \pi \), or \( w(T) \) of \( G' \) is at most one larger than the cost of \( G' \). Hence, \( w(T) \leq \text{cost}(G) \).

(ii) If \( A_2 \) has at least one \( T \)-vertex, then \( w(T) \) of \( G \) is three smaller than \( w(T) \) of \( G' \) since the \( T \)-vertex is connected with the three cc's of \( B_0 \). Thus, \( w(T) \leq \text{cost}(G) \).

(iii) Now suppose that the whole \( A_2 \)-vertices are in \( S \). However, \( w(T) \leq \text{cost}(G) \) nevertheless holds if \( B_1 \notin T \), since the number of cc's decreases at least by one. The same argument applied for each \( \mathcal{G} \) implies the lemma. □

Lemma 7. Suppose that \( \pi \) is a vertex partition such that each \( \mathcal{G}_i \) has at least one vertex not in \( T \), \( A_i \subseteq S \) and \( B_{i-1} \subseteq T \) for all \( i \geq 1 \). Then \( w_\pi(T) \leq \text{cost}_\pi(G) \) nevertheless hold if (i) both of the two vertices \( u_{i,0} \) and \( u_{i,0} \) of \( \mathcal{G}_i \) are in \( R \) or in \( S \) for some \( 1 \leq i \leq n \), or (ii) all of the three vertices \( x_{i,1,0} \), \( x_{i,2,0} \) and \( x_{i,3,0} \) of \( \mathcal{G}_{n+j} \) are in \( R \) or in \( S \) for some \( 1 \leq j \leq m \).

Proof. Fix some \( k \) and suppose that the condition (i) and (ii) is met for that \( k \). As shown before, \( w_\pi(\mathcal{G}) \leq \text{cost}_\pi(\mathcal{G}) \) for \( i \geq 2 \) and \( w_\pi(\mathcal{G}) \leq \text{cost}_\pi(\mathcal{G}) + 1 \). Therefore,
\[ \sum_{i \neq k} w_\pi(T_i) \leq \sum_{i \neq k} \text{cost}_\pi(G_i) + 1, \text{ i.e., } w(T) \text{ of } G' - \mathcal{G}_k \text{ is at most one larger than the cost of } G' - \mathcal{G}_k. \text{ (Recall that } G' \text{ is the collection of } \mathcal{G}_i \text{'s but no edges among them.) Then if } w_\pi(T_k) \leq \text{cost}_\pi(G_k) - 1, \text{ then } \sum_{i \neq k} w_\pi(T_i) + w_\pi(T_k) \leq \sum_{i \neq k} \text{cost}_\pi(G_i) + \text{cost}_\pi(G_k), \text{ which means } w(T) \leq \text{cost}(G'). \text{ Then we can conclude } w(T) \leq \text{cost}(G) \text{ by Lemma 2.}

Let us show that \( w_\pi(T_k) \leq \text{cost}_\pi(G_k) - 1 \). We only discuss the condition (ii). (i) is similar. So, let \( k = n + j \). First, suppose that all of the three vertices \( x_{j,1,0}, x_{j,2,0} \) and \( x_{j,3,0} \) in some \( \mathcal{G}_{n+j} \) are in \( S \). The following two cases on the partition of the other six vertices \( x_{j,1,1}, x_{j,1,2}, x_{j,2,1}, x_{j,2,2}, x_{j,3,1}, \) and \( x_{j,3,2} \) should be considered:

1. All of the six vertices are in \( T \). Then \( \text{cost}(\mathcal{G}_{n+j}) = 6 \) but \( w(T_{n+j}) = 5 \). Hence, \( w_\pi(T_k) \leq \text{cost}_\pi(G_k) - 1 \).
2. Moving some \( T \)-vertices into \( S \) or into \( R \) can only increase the cost.

Next, suppose that all of the three vertices \( x_{j,1,0}, x_{j,2,0} \) and \( x_{j,3,0} \) are in \( R \). There are six more cases: (3) All of the other six vertices are in \( T \). Then \( \text{cost}(\mathcal{G}_{n+j}) = 6 \) but \( w(T_{n+j}) = 5 \). (4) One or more of them is in \( S \) and the other five or less in \( T \). Easy and omitted. (5) One of them is in \( R \) and the others in \( T \). Then again \( \text{cost}(\mathcal{G}_{n+j}) = 6 \) but \( w(T_{n+j}) = 5 \). (6) Two of them are in \( R \) and the others in \( T \). If, for example, \( x_{j,1,1} \) and \( x_{j,2,1} \) are both in \( R \), then \( \text{cost}(\mathcal{G}_{n+j}) = 6 \) but \( w(T_{n+j}) = 5 \). As another example, if \( x_{j,1,1} \) and \( x_{j,1,2} \) are in \( R \), then the cost gets smaller by one and the number of \( \text{cc} \)’s also gets smaller by one. (7) Three or more of them in \( R \) and the others in \( T \). Omitted. (8) Some vertices are in \( S \) and the others in \( R \). Then the cost is equal to \( |S| \) but there are no new \( \text{cc} \)’s.

Now we shall consider the case where neither (i) nor (ii) in Lemma 7 holds which will completely exhaust all the cases.

**Lemma 8.** Suppose that \( \pi \) is a vertex partition such that each \( \mathcal{G}_i \) has at least one vertex not in \( T \). Also suppose that (i) \( A_i \subset S \) and \( B_{i-1} \subset T \) for all \( i \geq 1 \), (ii) at least one of the two vertices \( u_{i,0} \) and \( \overline{u}_{i,0} \) of \( \mathcal{G}_i \) is in \( T \) for all \( 1 \leq i \leq n \), and (iii) at least one of the three vertices \( x_{j,1,0}, x_{j,2,0} \) and \( x_{j,3,0} \) of \( \mathcal{G}_{n+j} \) is in \( T \) for all \( 1 \leq j \leq m \). Then \( \text{w}_\pi(T) \leq \text{cost}_\pi(G) \) nevertheless hold if the predicate \( f \) is not satisfiable.

**Proof.** It should be noted that the conditions (i)–(iii) are essentially the same as those mentioned in the proof of Lemma 1. Namely, (i) all \( A_i \)’s are in \( S \) and all \( B_i \)’s in \( T \), (ii) if \( u_i = \text{false} \) then \( u_{i,0} \) is in \( T \); otherwise, \( \overline{u}_{i,0} \) in \( T \), and (iii) the \( T \)-vertex corresponds to the \textit{true}-literal under the truth assignment. Since \( f \) is not satisfiable, there must be at least one \( T \)-\( T \)-edge, which connects one of the \( T \)-vertices (i.e., corresponding to the \textit{false} literal) mentioned in (ii) with one of the \( T \)-vertices (i.e., corresponding to the \textit{true} literal) in (iii). This reduces the number of \( \text{cc} \)’s of \( G \) compared to \( G' \) by one. Lemma 2 guarantees that the cost cannot be reduced by introducing edges between subgraphs. Thus, \( \text{w}_\pi(T) \leq \text{cost}_\pi(G). \)
Finally, we consider the case where some subgraph $\mathcal{G}_i$ has only $T$-vertices. Recall that $G$ has at least one vertex not in $T$ from the definition.

**Lemma 9.** Let $\pi$ be a vertex partition such that all the vertices of some $\mathcal{G}_i$ are in $T$. Then $w_\pi(T) \leq \text{cost}_\pi(G)$.

**Proof.** There are three cases to be considered: (i) $\mathcal{G}_1$ has at least one vertex not in $T$ (we simply say that $\mathcal{G}_1$ is not $T$) and all the other $\mathcal{G}_i$ has only $T$-vertices ($\mathcal{G}_i$ is $T$). (ii) $\mathcal{G}_1$ is not $T$ and at least one other $\mathcal{G}_i$ is not $T$ either. (iii) $\mathcal{G}_i$ is $T$. In the following we only discuss (ii); (i) and (iii) are easier and omitted.

Suppose that some $\mathcal{G}_i$ is not $T$ and $\mathcal{G}_{i+1}$ is $T$ for $2 < i < n + m - 1$. (The case where $\mathcal{G}_2$ is $T$ will be described later as a special case.) Then we can prove that $w_\pi(\mathcal{F}_i \cup \mathcal{F}_{i+1}) \leq \text{cost}_\pi(\mathcal{G}_i \cup \mathcal{G}_{i+1})$ for any $\pi$ by analyzing the following three cases:

1. $B_{i-1}$ in $\mathcal{G}_i$ has at least one $T$-vertex. Since the cc $\mathcal{G}_{i+1}$ connects to this $T$-vertex, $w_\pi(\mathcal{F}_i \cup \mathcal{F}_{i+1}) \leq \text{cost}_\pi(\mathcal{G}_i \cup \mathcal{G}_{i+1})$ follows from $w_\pi(\mathcal{F}_i) \leq \text{cost}_\pi(\mathcal{G}_i)$. (2) $B_{i-1}$ has at least one $R$-vertex. Three $T$-$R$-edges between $\mathcal{G}_i$ and $\mathcal{G}_{i+1}$ increase the cost by one.

Let $\sigma_1, \ldots, \sigma_k$ be the number of $\mathcal{G}_j$'s being $T$ preceded by $\mathcal{G}_i$ being not $T$. As shown above, the number of cc's in each group cannot surpass the cost. (If it consists of a single non-$T$ $\mathcal{G}_i$, then either Lemma 3 (for $2 < i < n$) or Lemma 4 (for $n + 1 < i < n + m$) applies.)

Thus, if $w(T) > \text{cost}(G)$ could hold, it must be for a vertex partition $\pi$ such that $w_\pi(\mathcal{F}_1) = \text{cost}_\pi(\mathcal{F}_1) + 1$. However, as shown in Lemma 6, this does not directly imply $w_\pi(T) > \text{cost}_\pi(G)$, namely, to be so, at least all $A_i$'s must be in $S$. Since we are now assuming that at least one $\mathcal{G}_i$ is $T$, this cannot happen.

Finally, we consider the special case; suppose that $\mathcal{G}_2$ is $T$. Then we can prove that $w_\pi(\mathcal{F}_1 \cup \mathcal{F}_2) \leq \text{cost}_\pi(\mathcal{G}_1 \cup \mathcal{G}_2)$ for any partition $\pi$: (1) All of the three vertices in $B_0$ are in $T$. Since these three $T$-vertices connect to the cc $\mathcal{G}_2$ and the number of cc's decreases by two, $w_\pi(\mathcal{F}_1 \cup \mathcal{F}_2) \leq \text{cost}_\pi(\mathcal{G}_1 \cup \mathcal{G}_2)$ follows from Lemma 5-(ii). (3) If $B_0$ has at least one vertex in $R$, then three $T$-$R$-edges between $\mathcal{G}_1$ and $\mathcal{G}_2$ increase the cost by one. $w_\pi(\mathcal{F}_1 \cup \mathcal{F}_2) \leq \text{cost}_\pi(\mathcal{G}_1 \cup \mathcal{G}_2)$ follows again from Lemma 5-(ii). (4) $B_0$ has only $S$-vertices. We can prove that $w_\pi(\mathcal{F}_1) \leq \text{cost}_\pi(\mathcal{G}_1) - 1$ for any $\pi$, which concludes that $w_\pi(\mathcal{F}_1 \cup \mathcal{F}_2) \leq \text{cost}_\pi(\mathcal{G}_1 \cup \mathcal{G}_2)$. Then, it is straightforward to show that $w_\pi(\mathcal{F}_1 \cup \mathcal{F}_2 \cup \cdots \cup \mathcal{F}_i) \leq \text{cost}_\pi(\mathcal{G}_1 \cup \mathcal{G}_2 \cup \cdots \cup \mathcal{G}_i)$ for $\mathcal{G}_2, \ldots, \mathcal{G}_i$ being $T$ but $\mathcal{G}_{i+1}$ being not $T$. Thus, the number of cc's in such a group also cannot surpass the cost, which concludes that $w_\pi(T) \leq \text{cost}_\pi(G)$. \[\square\]

One can easily verify that the degree of each vertex of $G$ is at most 15.
5. Proof of Theorem 2

Since $NS2F$ and $N1T$ are both in NP, $NS2F - N1T$ is in $D^P$. To show the completeness, we reduce the following problem, SAT-UNSAT which is known to be $D^P$-complete [14], to this problem: Given two 3-CNF predicates $f$ and $\tilde{f}$ over disjoint variable sets, the problem asks whether it is the case $f$ is satisfiable and $\tilde{f}$ is unsatisfiable. We can assume, without loss of generality, that both $f$ and $\tilde{f}$ use $n$ variables and contain $m$ clauses. For these predicates $f$ and $\tilde{f}$, we construct the following graph $G$: $G$ consists of five subgraphs $SG_1$, $\tilde{SG}_1$, $SG_2$, $\tilde{SG}_2$ and $SG_3$. (1) $SG_3$ is exactly the same as $SG_3$ in the proof of Theorem 1 but it consists of $2(n + m)$ $A_i$'s and $2(n + m)$ $B_i$'s. Note that only $B_0$ has three vertices. (2) $SG_1$ is associated with the variables of $f$. Again it is exactly the same as $SG_1$ of Theorem 1. (3) $SG_2$ is associated with the clauses of $f$ and is exactly the same as before. (4) $\tilde{SG}_1$ is associated with the variables of $\tilde{f}$ and consists of $n$ components $\tilde{G}_{u_i}$, $i = 1, \ldots, n$. $\tilde{G}_{u_i}$ is shown in Fig. 12. (5) $\tilde{SG}_2$ is associated with the clauses of $\tilde{f}$ and consists of $m$ components $\tilde{G}_{c_j}$, $j = 1, \ldots, m$. $\tilde{G}_{c_j}$ is shown in Fig. 13. (6) Edges between $SG_1$ and $SG_2$, $SG_1$ and $SG_2$, and $SG_3$ and $SG_1 \cup SG_2 \cup \tilde{SG}_1 \cup \tilde{SG}_2$ are also drawn similarly as before.

The key idea is as follows: The graph shown in Fig. 13 is similar to the graph in Fig. 5, but outer two vertices are missing and each degree of $\tilde{x}_{1,2,0}$ and $\tilde{x}_{1,3,0}$ is reduced by one. Suppose that the inner three vertices $\tilde{x}_{1,1,0}$, $\tilde{x}_{1,2,0}$ and $\tilde{x}_{1,3,0}$ of $\tilde{G}_{u_i}$ are in $S$ and the other four vertices in $T$. Then the number of cc's and the cost are both three. However, if the inner three vertices are in $R$ and the others in $T$, then the number of cc's is also three but the cost is reduced to two. This is the benefit of reducing the degree of the $R$-vertices, $\tilde{x}_{1,2,0}$ and $\tilde{x}_{1,3,0}$. (In the case of the graph shown in Fig. 5, even if the inner three vertices are in $R$, the cost remains three.) Thus, even if the old graph does not satisfy the condition of $NS2F$, the new graph may satisfy it. This is also true for $\tilde{SG}_1$ shown in Fig. 12.

Now suppose that $f$ is satisfiable and $\tilde{f}$ is not. We have to show that $G$ is in $NS2F - N1T$. (i) For $SG_3$, all $A_i$'s and all $B_i$'s are taken as $S$ and as $T$, respectively. Hence, we can separate $2 \times 2(n + m) + 1$ at the cost of $3 \times 2(n + m)$. (ii) For $SG_1$ and $SG_2$, we simulate the previous way of taking vertices as $S$ (recall that $f$ is satisfiable). The total cost is $2m + n$ and we can separate $3m + 2n$ components in total. (iii) From each component of $\tilde{SG}_1$, we take the inner two vertices as $R$ and the outer two vertices as $T$. The cost is $\lfloor 2/2 \rfloor = 1$. (iv) From each component of $\tilde{SG}_2$, we take $\tilde{x}_{j,1,0}$, $\tilde{x}_{j,2,0}$ and $\tilde{x}_{j,3,0}$ as $R$ and the others in $T$. The cost is $\lfloor (2 + 1 + 1)/2 \rfloor = 2$. Thus, we can separate two new cc's at the cost of one in (iii) and three ones at the cost of two in (iv). Now by a simple calculation, we can show that the number of cc's is one larger than the total cost. Thus $G$ is in $NS2F$.

Now let us show that $G$ is not in $N1T$. As before, we divide the entire graph $G$ into $2(n + m)$ subgraphs, $G_1, \ldots, G_{n+m}$ and $\tilde{G}_1, \ldots, \tilde{G}_{n+m}$, where each $G_i$ has $A_i$ and $B_{i-1}$ and each $\tilde{G}_i$ has $A_i$ and $B_{i-1}$. The following two claims are very similar to Claim 1 in the proof of Lemma 3:
Claim 2. At most $h$ cc's can be separated from $\tilde{G}_{n_i}$ at the cost of $h - 1$ for any vertex partition. Same for $\tilde{G}_{n_i}$ for $2 \leq i \leq n$.

Proof. There are the following six cases: (1) All of the four vertices in $\tilde{G}_{n_1}$ are in $T$. We obtain one cc at no cost. (2) If there exists at least one $S$-vertex, then the cost is also at least one. Since the number of cc's is at most two, this claim holds. (3) One of the four vertices is in $R$ and the others in $T$. If one of $\tilde{u}_{1,0}$ and $\tilde{u}_{1,1}$ is in $R$, then two cc's can be separated at the cost of one; otherwise, one cc at no cost. (4) Two of them are in $R$. Then at most two cc's can be separated at the cost of one. This is the benefit of reducing the degree of the $R$-vertices as described before. (5) If three of them are in $R$, there is only one cc. (6) All of them are in $R$. There is no cc. □

Claim 3. At most $h$ cc's can be separated from $\tilde{G}_{c_j}$ at the cost of $h - 1$ for any vertex partition. Same for $\tilde{G}_{c_j}$ for $2 \leq j \leq m$.

Proof. The number of cc's increases by at most one from Claim 2. One can see that $\tilde{G}_{c_j}$ consists of the two subgraphs, one consists of four vertices $\tilde{x}_{1,2,0}$, $\tilde{x}_{1,2,1}$, $\tilde{x}_{1,3,0}$ and $\tilde{x}_{1,3,1}$, which is the same as $\tilde{G}_{u_1}$, and the other is a complete graph with three vertices, $\tilde{x}_{1,1,0}$, $\tilde{x}_{1,1,1}$ and $\tilde{x}_{1,1,2}$, denoted by, say, $H$. If $H$ has at least one $R$- or one $S$-vertex, then the claim holds by Claim 2 since the cost increases. Now suppose that all vertices of $H$ are in $T$, and consider the two vertices $\tilde{x}_{1,2,0}$ and $\tilde{x}_{1,3,0}$ that are connected to $H$. If (at least) one of them is in $T$, then the number of cc's does not increase compared to Claim 2. If both are in $R$, then the cost increases. So, the remaining cases are (i) one is in $S$ (and the other in $R$) and (ii) both are in $S$. (i) Recall that we are now assuming that two outer vertices $\tilde{x}_{1,2,1}$ and $\tilde{x}_{1,3,1}$ are both in $T$. So the cost is two ($= \left\lceil \frac{3}{2} \right\rceil + 1$) and the number of cc's is three. (ii) The number of cc's is at most three and its cost is at least two. □

Lemma 10. Let $\pi$ be a vertex partition such that $\tilde{G}_1$ has at least one vertex not in $T$. Then, $w_\pi(\tilde{G}_1) \leq \text{cost}_A(\tilde{G}_1)$. Same for $\tilde{G}_i$ for $2 \leq i \leq n + m$. 
**Proof.** Note that each $\tilde{A}_i \cup \tilde{B}_{i-1}$ are the same as $A_2 \cup B_1$ in the proof of Theorem 1, i.e., every $\tilde{B}_{i-1}$ in $\tilde{G}_i$ has only two vertices. Considering the similar five cases as the proof of Lemma 3, this lemma follows from Claims 2 and 3. \qed

Since we are now discussing $N1T$, the vertex partition only includes $S$ and $T$. Then, intuitively, we can no longer take the advantage of the benefit mentioned before. We can use almost the same approach as Lemmas 6 and 7 to prove that $w(T) = |S|$ for any partition such that (i) both of $\tilde{u}_{i,0}$ and $\tilde{u}_{i,6}$ in $\tilde{G}_1$ are in $S$ for some $i$, or (ii) all of $\tilde{x}_{j,1,0}$, $\tilde{x}_{j,2,0}$ and $\tilde{x}_{j,3,0}$ in $\tilde{G}_2$ are in $S$ for some $j$. For the case where neither (i) nor (ii) holds, we can apply the same argument of Lemma 8. Thus, $G$ is not in $N1T$.

Next suppose that both $f$ and $\tilde{f}$ are satisfiable. Then it is straightforward to show that the graph $G$ is in $N1T$, namely, it is not in $NS2F - N1T$.

Finally, suppose that $f$ is not satisfiable. Recall that each $\tilde{B}_{i-1}$ in $\tilde{G}_i$ has only two vertices. A key observation is that whether or not $\tilde{f}$ is satisfiable, we can separate at most $h_1$ cc’s at the cost of $h_1$ from the subgraph $\tilde{G}_1 \cup \cdots \cup \tilde{G}_{n+m}$ by Lemma 10. Similarly, since $f$ is not satisfiable, as was shown in the proof of Theorem 1, we can separate at most $h_2$ cc’s at the cost of $h_2$ from the subgraph $\tilde{G}_1 \cup \cdots \cup \tilde{G}_{n+m}$ even if $B_0$ has three vertices. That means $G$ is not in $NS2F$ and, hence, not in $NS2F - N1T$. One can check that the degree of each vertex is at most 15. \qed

6. **Proof of Theorem 4**

We show that the problem of deciding, given graph $G=(V,E)$, whether or not there is a subset $S$ of $V$ such that $t \cdot w(G - S) > |S|$ is NP-complete. Since we can check the condition in polynomial time for particular $S$, the problem is in NP. To show the NP-hardness of $N1T$, 3-SAT is reduced to $N1T$ again. Let $t = a/b$. Then the condition for $N1T$ can be written as $(w(G - S))/b > |S|/a$. So, if we let $|S| = a \cdot h$ then the inequality becomes $w(G - S) > b \cdot h$. That means we have to create at least $b \cdot h + 1$ cc’s at the cost (the number of vertices taken as $S$) of $a \cdot h$ for some $h$ to satisfy the inequality.

The reduction is similar to before: Again the graph $G$ that is reduced from a predicate $f$ consists of three subgraphs $SG_1$, $SG_2$ and $SG_3$. Each component $G_u$ of $SG_1$ looks like Fig. 14. Namely, $u_{i,0}(u_{i,6})$ of Fig. 4 is replaced by $U_{i,0}(U_{i,6})$ that is a complete graph of $a$ vertices denoted by $K_a$. Also $u_{i,1}$ and $u_{i,2}(u_{i,1}$ and $u_{i,2}$) are replaced by $b$ independent vertices denoted by $U_{i,1}(U_{i,1})$. We again use the complete bipartite connection among them. The construction is similar for each component $G_v$ of $SG_2$ as shown in Fig. 15. Edges between $SG_1$ and $SG_2$ are drawn similarly as before. For $SG_3$, each $A_i$ is $K_{3a}$ and each $B_j$ is independent $(3b - 1)$ vertices. Only $B_0$ contains independent $3b$ vertices. As before, we introduce the complete bipartite connection between $A_1$ and $B_0$, $B_0$ and $A_2$, and so on. Also, the complete bipartite connections exist between the $3a$ vertices of $A_i$ and the $2a + 2b$ vertices of $G_{i_0}$ for each $i = 1, \ldots, n$, and between $A_{n+j}$ and the $3a + 3b$ vertices of $G_v$ for each $j = 1, \ldots, m$. 


We first show that if an instance $f$ is satisfiable, then $G$ is in $N_{itT}$ ($t = a/b$). Recall the proof of Theorem 1. For example, we took $u_{i,0}$ from $SG_1$ and obtained two cc's. This time we take $U_{i,0}$ (i.e., $a \times 1$ vertices) and obtain $b \times 1 + 1$ cc's (in this case, $h = 1$). As for $SG_2$, by taking two $K_a$'s ($a \times 2$ vertices) out of the three $K_a$'s, we can obtain $b \times 2 + 1$ cc's. Which one of the three $K_a$'s is not taken is determined by the truth assignment in the same way as before. Thus we can separate $b(n + 2m) + (n + m)$ cc's by taking $a(n + 2m)$ vertices (here $h = n + 2m$) from $SG_i$ and $SG_2$. As for $SG_3$, if we take all $A_i$'s ($3a(n + m)$ vertices), we can obtain $(3b - 1)(n + m) + 1$ cc's. In total we obtain $b(4n + 5m) + 1$ cc's by taking $a(4n + 5m)$ vertices. Thus one can see that the graph $G$ is $N_{itT}$.

It remains to prove that if $f$ is not satisfiable, then the condition for $N_{itT}$ cannot be satisfied for any subset $S$. Suppose for contradiction that the graph $G$ is in $N_{itT}$, namely, the condition is met for some $S$. The first observation is that we have to take all the vertices of $U_{i,0}$ whenever we take some vertices from $U_{i,0}$.

**Lemma 11.** If the condition for $N_{itT}$ can be satisfied by taking some but not all vertices in $U_{i,0}$, then it can be also satisfied by taking no vertices in $U_{i,0}$. Similarly, for $\overline{U_{i,0}}$, $X_{j,1,0}$, $X_{j,2,0}$, $X_{j,3,0}$ and $A_k$.

**Proof.** Note that $U_{i,0}$ is a complete graph and has complete bipartite connections with adjacent vertices. Hence, taking not the whole but a part of $U_{i,0}$ does not create
any new cc but only increases $|S|$. In other words, if we do not take those vertices, the number of cc’s does not decrease and $|S|$ does decrease, which claims the lemma.

Let $G_0$ be the graph obtained from $G$ by removing all edges between $SG_1$ and $SG_2$. Note that for any $S \subseteq V$, the number of cc’s obtained from $G$ by taking $S$ ($= w(G - S)$) is at most the number of cc’s obtained from $G_0$ by taking the same $S$ ($= w(G_0 - S)$). For a while, we take a look at this $G_0$. Remember that each $G_u$ (similarly for each $G_c$) is connected to $A_k$ by complete bipartite connection. Hence, we cannot create any new cc from $G_u$ unless all vertices in $A_k$ are taken as $S$. So, suppose that all vertices in $A_k$ are taken. Then we only have to consider very few different ways of taking vertices in $G_u$ and $G_c$, each of which is now isolated from $G_0$:

1. If the condition for $N(tT$ can be satisfied, it can be done so by one of the following three ways of taking vertices as $S$ for $G_u$: (The reason is obvious by Lemma 11 and by the fact that taking some of the independent vertices does not create any new cc.)
   (1-1) Take all $a$ vertices in $U_{l,0}$ or $U_{l,0}$ and we can get $b + 1$ cc’s from $G_0$. (1-2) Take $2a$ vertices in $U_{l,0}$ and $U_{l,0}$ and we can get $2b$ cc’s. (1-3) Take no vertices and we can get one cc ($G_u$ itself).

2. Similarly, for $G_c$: (2-1) Take one $K_a$ (i.e., take $a$ vertices and we get $b + 1$ cc’s), (2-2) two $K_a$’s ($2a$ vertices, $2b + 1$ cc’s), (2-3) three $K_a$’s ($3a$ vertices, $3b$ cc’s) and (2-4) zero vertices and one cc ($G_c$ itself).

We next show that the condition can be satisfied only if all $n + m A_k$’s in $SG_3$ are taken. Suppose otherwise that $d A_k$’s are taken ($1 \leq d \leq n + m - 1$). Then, since at least one $A_k$ remains untaken, the number of $B_k$’s which are cut off from the neighbouring $A_k$’s is at most $d - 1$. That means the number of cc’s obtained from $SG_3$ is at most $(d - 1)(3b - 1) + 1$. (“+1” comes from $B_0$ that contains one vertex more than other $B_k$.)

If some $A_k$ is taken, then we can create further new cc’s from its neighbouring $G_u$ or $G_c$. However, what we can do best is to obtain $b \cdot h + 1$ cc’s by taking $a \cdot h$ vertices ($h = 0, 1$ or 2) as shown in above (1) and (2). In total, we can obtain

$$(d - 1)(3b - 1) + 1 + \sum_{l=1}^{d} (b \cdot h_{k_l} + 1) = b \left( 3d + \sum_{l=1}^{d} h_{k_l} \right) - 3b + 2$$

cc’s by taking

$$3a \cdot d + \sum_{l=1}^{d} a \cdot h_{k_l} = a \left( 3d + \sum_{l=1}^{d} h_{k_l} \right)$$

vertices. This does not satisfy the condition. Thus we can conclude that we have to take all $A_k$’s.

If we take all $A_k$’s, then the number of cc’s obtained from $SG_3$ increases to $(n + m)(3b - 1) + 1$. By a similar calculation as above, the total number of cc’s
including those obtained from $SG_1$ and $SG_2$ is

$$
(n + m)(3b - 1) + 1 + \sum_{k=1}^{n+m} (b \cdot h_k + 1) = b \left( 3n + 3m + \sum_{k=1}^{n+m} h_k \right) + 1
$$

and the number of the vertices taken is

$$
3a(n + m) + \sum_{k=1}^{n+m} a \cdot h_k = a \left( 3n + 3m + \sum_{k=1}^{n+m} h_k \right).
$$

Thus the condition is barely met. It should be noticed that one can only use the ways $(l-1)$ or $(l-3)$, and $(2-1)$, $(2-2)$ or $(2-4)$ of taking vertices where "$+_1$" in the number of cc's plays an important role. In other words, at least one of $U_{i,0}$ and $U_{i,0}$ in $G_u$ and at least one of $X_j,1,0$, $X_j,2,0$ and $X_j,3,0$ in $G_c$ cannot be taken, which we call remaining $K_a$'s.

Now remember that there are edges between $SG_1$ and $SG_2$ in $G$ that were ignored in $G_0$. Suppose that there are such edges between some remaining $K_a$'s. Then the condition is no longer met since those two $K_a$'s can be counted as only one cc although they were counted as two cc's in the above calculation. Thus, we can conclude that there are no edges between any two remaining $K_a$'s if the condition is met. Now one can see that we can construct a truth assignment that makes $f$ true in the same way as described in the proof of Theorem 1. This however contradicts to the assumption that $f$ is not satisfiable. One can check that the degree of each vertex of $G$ is at most $6a + 9b - 3$. □

References


