6 February 2002

# Generalizations of Matched CNF Formulas

**(Extended Abstract)**

Stefan Szeider

*Institute of Discrete Mathematics, Austrian Academy of Sciences, Sonnenfelsgasse 19, A–1010 Vienna, Austria, stefan.szeider@oeaw.ac.at*

**Abstract**

A CNF formula is "matched" if its associated bipartite graph (whose vertices are clauses and variables) has a matching which covers all clauses. Matched CNF formulas are always satisfiable, and can be recognized efficiently by matching algorithms. We generalize this concept and cover clauses by collections of bicliques (complete bipartite graphs). It turns out that such generalizations indeed give raise to larger classes of satisfiable CNF formulas (which we term "biclique satisfiable"). We show, however, that recognition of biclique satisfiable CNF formulas is NP-complete, and remains NP-hard if the size of bicliques is bounded by a fixed integer.

We call a CNF formula "graph satisfiable" if satisfiability is determined solely by its associated bipartite graph. Matched and biclique satisfiable formulas are graph satisfiable; hence, graph satisfiable CNF formulas can be viewed as the best possible generalization of matched CNF formulas. We pinpoint the exact complexity of recognizing graph satisfiable CNF formulas by showing $\Pi_2^{\mathrm{P}}$-completeness.

*Key words:* SAT problem, bipartite graph, matched formula, biclique cover, NP-completeness, $\Pi_2^{\mathrm{P}}$-completeness, polynomial hierarchy.

## 1 Introduction

The *satisfiability problem* (SAT) asks whether a given propositional formula $F$ in conjunctive normal form (CNF) has a satisfying truth assignment (i.e., $F$ is satisfiable). SAT is the first problem which has been shown to be NP-complete. Since this problem is important in many areas (e.g., Artificial Intelligence and Operations Research), considerable efforts have been made to identify classes of satisfiable CNF formulas which can be recognized in polynomial time; see [5,8] for surveys.

Our point of departure is a certain class of satisfiable CNF formulas which

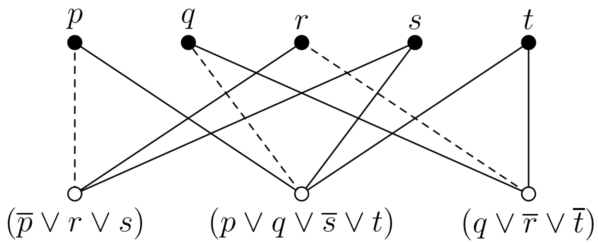


Fig. 1. Graph of a matched CNF formula.

can be recognized in polynomial time by matching algorithms. To define this class, we associate to a CNF formula $F$ a bipartite graph $G(F)$ where one class of vertices consists of the variables occurring in $F$, and the other class consists of the clauses of $F$; a variable $v$ and a clause $C$ are joined by an edge if $v$ occurs in $C$ (positively or negatively). For example, Figure 1 exhibits the graph associated with the formula

$$F = (\overline{p} \vee r \vee s) \wedge (p \vee q \vee \overline{s} \vee t) \wedge (q \vee \overline{r} \vee \overline{t}). \tag{1}$$

(The meaning of dashed lines will be explained later.) Note that polarity of literals in $F$ is not taken into account by forming $G(F)$.

In [14] the following simple but nevertheless fruitful observation was made (see also [1]).

> *If $G(F)$ has a matching $M$ (i.e., a set of independent edges) such that every clause of $F$ is incident with some edge in $M$, then $F$ is satisfiable.*

For, we can obtain a satisfying truth assignment by setting a variable TRUE if it occurs positively in the clause to which it is matched, and FALSE otherwise (the value of variables which are not matched to any clause is arbitrary).

Following [5] we call a CNF formula *matched* if the associated graph $G(F)$ has a matching which covers all clauses. For example, the graph in Figure 1 has such matching (indicated by dashed lines); thus the formula in (1) is satisfiable. Since a matching of maximum size can be found in polynomial time (see, e.g., [11]), it can be decided efficiently whether a given CNF formula is matched.

The concept of matched CNF formulas is used in [4] to obtain polynomial-time algorithms for recognition of minimal unsatisfiable CNF formulas (i.e., unsatisfiable CNF formulas which become satisfiable by removing any clause) with fixed clause-variable difference. Moreover, this concept provides the basis for a powerful class of autark assignments ("matching autarkies"), studied in [9,10].

It would be desirable to find suitable generalizations of matched CNF formulas to extend the quoted applications.

In Section 2 we generalize matched CNF formulas, covering the associated graph by disjoint bicliques (complete bipartite graphs). This results in a class of satisfiable CNF formulas ("biclique satisfiable CNF formulas") which properly contains all matched CNF formulas (a matching is nothing but a collection of disjoint bicliques on two vertices each). We show, however, that recognition of biclique satisfiable CNF formulas is NP-complete (and remains NP-hard if the size of bicliques is bounded from above). We present a reduction from graph 3-colorability in Appendix A.

The question arises whether "covering by bicliques" is a too specific concept, since a more general approach might yield a larger class of satisfiable CNF formulas which can be recognized efficiently. In order to tackle this question, we consider in Section 3 such bipartite graphs $G$ for which *every* CNF formula $F$ with $G(F) = G$ is satisfiable. The corresponding class of CNF formulas contains exactly those which are satisfiable and remain satisfiable under arbitrary replacement of literals by their complement. We call such CNF formulas "graph satisfiable." Clearly if a CNF formula is biclique satisfiable (or matched), then it is also graph satisfiable, but the converse does not hold. It turns out, however, that graph satisfiable CNF formulas are even harder to recognize than biclique satisfiable CNF formulas. We pinpoint the exact complexity of graph satisfiability, and show that recognition is $\Pi_2^{\mathrm{P}}$-complete ($\Pi_2^{\mathrm{P}}$ is a complexity class in the second level of the polynomial hierarchy, see [12,13]; some authors use $\Pi_2\mathrm{P}$ or co-$\mathrm{NP}^{\mathrm{NP}}$ to denote this complexity class). For the $\Pi_2^{\mathrm{P}}$-completeness proof we present a reduction of an enhanced graph coloring problem, where it is asked whether every 2-coloring of a specified set of vertices of a given graph can be extended to a proper 3-coloring of all vertices. This reduction arises from a modification of the construction in Appendix A, and is presented in Appendix B.

The following diagram visualizes our main results.

| matched | $\subsetneq$ | biclique satisfiable | $\subsetneq$ | graph satisfiable |
|---|---|---|---|---|
| (polynomial time) | | (NP-complete) | | ($\Pi_2^{\mathrm{P}}$-complete) |

Both completeness results hold even for 3-*CNF formulas* (formulas with at most three literals per clause).

## 2 Covering by bicliques

A single edge can be considered as a biclique (i.e., a complete bipartite graph) on two vertices, and so we ask if we can use a collection of bicliques instead of a matching to certify satisfiability. The following observation gives a reason for such approach.

**Lemma 1** *Let $F$ be a CNF formula on $n$ variables and $m$ clauses. If $G(F)$ is a biclique and $m < 2^n$, then $F$ is satisfiable.*

**PROOF.** Every clause of $F$ has exactly $n$ literals; hence every clause of $F$ discards exactly one of the $2^n$ truth assignments. □

For example, the CNF formula $(p \vee q) \wedge (p \vee \overline{q}) \wedge (\overline{p} \vee q)$ is satisfiable by means of Lemma 1 (however, it is not matched).

The next idea is that we can certify satisfiability by using several bicliques to cover all clauses in $G(F)$, as long as every variable belongs to at most one biclique.

**Lemma 2** *A CNF formula $F$ is satisfiable if there is a collection $\mathcal{C} = \{X_1, \ldots, X_k\}$ of sets of vertices of $G(F)$ such that*

(i) *every clause of $F$ belongs to some $X_i$, $1 \leq i \leq k$,*
(ii) *every variable of $F$ belongs to at most one $X_i$, $1 \leq i \leq k$,*
(iii) *$X_i$ induces in $G(F)$ a biclique on $n_i$ variables and $m_i < 2^{n_i}$ clauses, $1 \leq i \leq k$.*

**PROOF.** By means of Lemma 1, we find for every $i = 1, \ldots, k$ some truth assignment $t_i$ to the variables in $X_i$ such that all clauses in $X_i$ are satisfied. Since every variable of $F$ belongs to at most one $X_i$, we can assemble $t_1, \ldots, t_k$ consistently to a satisfying truth assignments of $F$. □

We call a CNF formula *biclique satisfiable* if it fulfills the conditions stated in Lemma 2. See Figure 2 for an example, where bicliques are indicated by different types of dashed lines. The question arises whether biclique satisfi-

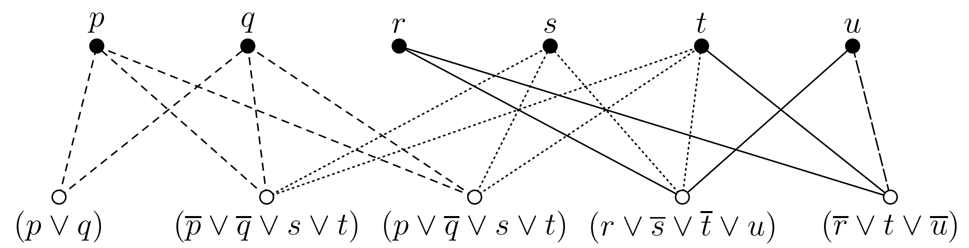


Fig. 2. Graph of a biclique satisfiable CNF formula.

able CNF formulas can be recognized efficiently. Unfortunately, our results of Appendix A show the following.

**Theorem 1** *Recognition of biclique satisfiable CNF formulas is* NP*-complete, and remains* NP*-hard for 3-CNF formulas.*

Are there any restrictions of the general concept of biclique satisfiability which admit efficient recognition? The following restrictions seem to be natural candidates for this approach.

- A CNF formula $F$ is *disjointly biclique satisfiable* if there is a collection $\mathcal{C}$ which satisfies the conditions of Lemma 2, and the elements of $\mathcal{C}$ are mutually disjoint.
- Let $k$ be a positive integer. A CNF formula $F$ is *$k$-biclique satisfiable* if there is a collection $\mathcal{C}$ which satisfies the conditions of Lemma 2, and each $X_i$ contains at most $k$ variables.
- Let $k$ be a positive integer. A CNF formula $F$ is *disjointly $k$-biclique satisfiable* if there is a collection $\mathcal{C}$ which satisfies the conditions of Lemma 2, the elements of $\mathcal{C}$ are mutually disjoint, and each $X \in \mathcal{C}$ contains at most $k$ variables.

**Lemma 3** *For a CNF formula $F$ the following are equivalent.*

(i) *$F$ is matched;*
(ii) *$F$ is 1-biclique satisfiable;*
(iii) *$F$ is disjointly 1-biclique satisfiable.*

**PROOF.** Obviously, $F$ is disjointly 1-biclique satisfiable if and only if it is matched. Furthermore, if $\mathcal{C}$ certifies a CNF formula $F$ to be 1-biclique satisfiable, then there is a subset $\mathcal{C}'$ of $\mathcal{C}$ which certifies $F$ to be disjointly 1-biclique satisfiability. $\square$

Our results of Appendix A yield that even under the proposed restrictions, recognition of biclique satisfiable CNF formulas remains intractable:

**Theorem 2** *Recognition of disjointly biclique satisfiable CNF formulas is* NP *complete. Moreover, for any $k \geq 2$, recognition of $k$-biclique satisfiable CNF formulas, and recognition of disjointly $k$-biclique satisfiable CNF formulas is* NP*-complete.* NP*-hardness holds even for* 3*-CNF formulas.*

## 3 Graph satisfiability

We say that a CNF formula $F$ is *graph satisfiable* if every CNF formula $F'$ with $G(F') = G(F)$ is satisfiable. Clearly matched and biclique satisfiable CNF formulas are graph satisfiable, since polarity of literals is not taken into account in the associated bipartite graph. However, there are graph satisfiable CNF formulas which are not biclique satisfiable; take, for example,

$$F = (p \vee q \vee r) \wedge (p \vee q \vee r) \wedge (p \vee q \vee r) \wedge (p \vee q) \wedge (p \vee q).$$

(That $F$ is indeed graph satisfiable can be established by a similar argument as used in the proof of Lemma 1.)

The reduction from 3-colorability presented in Appendix A shows NP-hardness of graph satisfiability. For showing membership in NP, one could try to extend the concept of biclique satisfiability, allowing graphs which are not bicliques as elements of a cover. Such approach, however, cannot be successful unless P = NP, as our next result shows (for a proof see Appendix B).

**Theorem 3** *Recognition of graph satisfiable CNF formulas is* $\Pi_2^{\mathrm{P}}$*-complete, and remains* $\Pi_2^{\mathrm{P}}$*-hard for* 3*-CNF formulas.*

We recall that the complexity class $\Pi_2^{\mathrm{P}}$ belongs to the second level of the polynomial hierarchy, introduced in [13]; $\Pi_2^{\mathrm{P}}$ contains the complements of all properties which can be checked in nondeterministic polynomial time by an NP-oracle (or, equivalently, by a co-NP-oracle).

It is easy to see that graph satisfiability belongs to $\Pi_2^{\mathrm{P}}$: if a given CNF formula $F$ is not graph satisfiable, then we can guess an unsatisfiable CNF formula $F'$ with $G(F) = G(F')$; unsatisfiability of $F'$ can be checked by a co-NP-oracle. For showing $\Pi_2^{\mathrm{P}}$-completeness of graph satisfiability, we use a reduction from an enhanced 3-colorability problem (see Appendix B).


## Acknowledgment

The idea to generalize matchings by collections of bicliques emerged through discussions with Oliver Kullmann at his research visit in Vienna, August 2001.


## References


[1] R. Aharoni and N. Linial. Minimal non-two-colorable hypergraphs and minimal unsatisfiable formulas. *J. Combin. Theory Ser. A*, 43:196–204, 1986.

[2] R. Diestel. *Graph Theory*, volume 173 of *Graduate Texts in Mathematics*. Springer Verlag, 2nd edition, 2000.

[3] T. Eiter and G. Gottlob. Complexity results for some eigenvector problems. *International Journal of Computer Mathematics*, 76(1):59–74, 2000.

[4] H. Fleischner, O. Kullmann, and S. Szeider. Polynomial-time recognition of minimal unsatisfiable formulas with fixed clause-variable difference. To appear in *Theoretical Computer Science*.

[5] J. Franco and A. Van Gelder. A perspective on certain polynomial time solvable classes of satisfiability. To appear in *Discr. Appl. Math.*

[6] M. R. Garey and D. R. Johnson. *Computers and Intractability.* W. H. Freeman and Company, New York, 1979.

[7] P. Hell. Graph packings. *Electronic Notes in Discrete Mathematics*, 5, 2000.

[8] H. Kleine Büning and T. Lettman. *Propositional logic: deduction and algorithms.* Cambridge University Press, Cambridge, 1999.

[9] O. Kullmann. Lean clause-sets: Generalizations of minimally unsatisfiable clause-sets. Submitted to *Discr. Appl. Math.*

[10] O. Kullmann. Investigations on autark assignments. *Discr. Appl. Math.*, 107(1-3):99–137, 2000.

[11] L. Lovász and M. D. Plummer. *Matching Theory*, volume 29 of *Annals of Discrete Mathematics.* Elsevier Science Publishers, North-Holland, 1986.

[12] C. H. Papadimitriou. *Computational Complexity.* Addison-Wesley, 1994.

[13] L. J. Stockmeyer. The polynomial-time hierarchy. *Theoretical Computer Science*, 3(1):1–22, 1976.

[14] C. A. Tovey. A simplified NP-complete satisfiability problem. *Discr. Appl. Math.*, 8(1):85–89, 1984.

## Appendix A. NP-completeness of biclique satisfiability

We will obtain the quoted NP- and $\Pi_2^{\mathrm{P}}$-completeness results in a graph theoretic framework; therefore we fix some terminology. For graph theoretic terminology not defined here, the reader is referred to [2].

All graphs considered are finite and simple. We denote the vertex set and the edge set of a graph $G$ by $V(G)$ and $E(G)$, respectively. If $G$ is bipartite, then we assume that sets $V_1(G), V_2(G) \subseteq V(G)$ give a fixed bipartition.

Let $G$ be a bipartite graph. A *biclique cover* $\mathcal{C}$ of $G$ is a set $\mathcal{C} = \{X_1, \dots, X_r\}$ of subsets of $V(G)$ such that each $X_i$ induces in $G$ a biclique.

A vertex $v$ is *covered* by $\mathcal{C}$ if $v$ is contained in $\bigcup_{i=1}^{r} X_i$; otherwise $v$ is *exposed* by $\mathcal{C}$. A biclique cover is *total* if every vertex in $V_2(G)$ is covered by $\mathcal{C}$ (vertices in $V_1(G)$ may remain exposed).

Let $\mathcal{C}$ be a biclique cover of $G$ and $G'$ a subgraph of $G$. Then $\mathcal{C}$ induces a biclique cover $\mathcal{C} \upharpoonright G' := \{\ X \in \mathcal{C} \mid X \subseteq V(G')\ \}$ of $G'$. Observe that it is possible that $\mathcal{C}$ is a total biclique cover of $G$, but $\mathcal{C} \upharpoonright G'$ is not a total biclique cover of $G'$.

A biclique cover $\mathcal{C}$ of $G$ is called *admissible* if

(i) for every $X \in \mathcal{C}$ we have $|V_2(G) \cap X| < 2^{|V_1(G) \cap X|}$, and
(ii) for any distinct $X, X' \in \mathcal{C}$ we have $V_1(G) \cap X \cap X' = \emptyset$

In the sequel we focus on admissible biclique covers; hence we use the term "*cover*" as a shorthand for "admissible biclique cover."

Note that if $\mathcal{C} = \{X_1, \dots, X_r\}$ is a total cover of $G$, then setting

$$Y_i := X_i \setminus \bigcup_{j=1}^{i-1} (X_j \cap V_2(G))$$

defines a total cover $\mathcal{C}' := \{\ Y_i \mid V_2(G) \cap Y_i \neq \emptyset,\ 1 \leq i \leq r\ \}$; elements of $\mathcal{C}'$ are mutually disjoint.

A 3-*coloring* of a graph $G$ is a map $\chi : V(G) \to \{1, 2, 3\}$. A 3-coloring is *proper* if $\chi(w) \neq \chi(v)$ for every edge $vw \in E(G)$. Furthermore, $G$ is called 3-*colorable* if $G$ has some proper 3-coloring.

It is well-known that the problem whether a given graph is 3-colorable is NP-complete, see e.g., [6] (evidently the problem remains NP-complete if only graphs with minimum degree 2 are considered). We will reduce this problem to the problem whether a given bipartite graph has a total cover. We note that the latter problem has some similarity to "graph packing" problems studied elsewhere (see [7] for a survey), but is apparently not entailed by known results from this area.

For the reduction, we will use the following two types of gadgets.

A *color gadget* is a bipartite graph $H_c$ with $V_1(H_c) = \{x_1, x_2\}$ and $V_2(H_c) = \{y_1, y_2, y_3, w\}$ as depicted in Figure A.1. We call the vertices $y_i$ *outputs* of $H_c$; we will also use the notation $\mathsf{out}_i(H_c) := y_i$, $i = 1, 2, 3$.

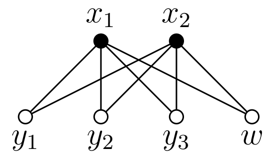


Fig. A.1. Color gadget $H_c$.

A color gadget $H_c$ has no total cover. However, for every $i \in \{1, 2, 3\}$ there is a (unique) cover $\mathcal{C}_i$ such that $y_i$ is the only vertex which remains exposed (e.g., $\mathcal{C}_1 = \{\{x_1, x_2, y_2, y_3, w\}\}$).

Let $q \geq 2$ be an integer. A *fan-out gadget* is a bipartite graph $H_f$ with $V_1(H_f) = \{\ a_i, b_i, z_i \mid i = 1, \ldots, q-1\ \} \cup \{z_q\}$ and $V_2(H_f) = \{\ c_i, d_i, e_i \mid i = 1, \ldots, q-1\ \}$, as depicted in Figure A.2(a) for $q = 2$ and in Figure A.2(b) for $q > 2$. We call $\mathsf{in}(H_f) := a_1$ the *input* of $H_f$, and the vertices in $\mathsf{out}(H_f) := \{z_1, \ldots, z_q\}$ the *outputs* of $H_f$.

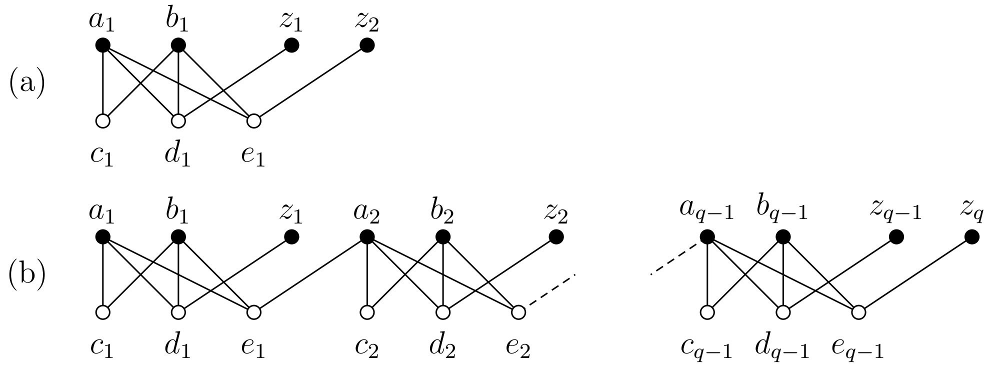


Fig. A.2. Fanout gadgets $H_f$.

**Lemma A.1** *A fan-out gadget $H_f$ has a total cover which exposes all outputs. There is exactly one total cover $\mathcal{C}$ of $H_f$ which exposes the input; such $\mathcal{C}$, however, covers all outputs.*

**PROOF.** Evidently $\{\ \{a_i, b_i, c_i, d_i, e_i\} \mid i = 1, \ldots, q-1\ \}$ is a total cover of $H_f$ which exposes all outputs. For the second part of the lemma, we construct $\mathcal{C}$

as follows. Since $c_1$ must be covered without using the input $a_1$, we have no choice but using $\{b_1, c_1\}$ to cover $c_1$. Whence the observation holds for $q = 2$, since we must use $\{z_1, d_1\}$ and $\{z_2, e_1\}$ to cover $d_1$ and $e_1$, respectively. For $q > 2$, we must use $\{e_1, a_2\}$ to cover $e_1$, and as above, $\{b_2, c_2\}$ to cover $c_2$, and so forth. The cover $\mathcal{C}$ defined by this process satisfies the claimed property. $\square$

Let $G$ be any finite graph with minimum degree 2. We are going to construct a bipartite graph $H_G$ from $G$. For each vertex $v$ of $G$ we will construct a graph $H(v)$; then we will put all these graphs together to form $H_G$.

Let $v \in V(G)$. We take a color gadget $H_{\mathrm{c}}(v)$, and fan-out gadgets $H_{\mathrm{f}}(v, j)$, $j = 1, 2, 3$ with $d_G(v)$ outputs each. For $j = 1, 2, 3$ we connect $H_{\mathrm{c}}(v)$ with $H_{\mathrm{f}}(v, j)$ by adding an edge between $\mathsf{out}_j(H_{\mathrm{c}}(v))$ and $\mathsf{in}(H_{\mathrm{f}}(v, j))$. We denote the obtained graph by $H(v)$.

Let

$$Z := \bigcup_{1 \le j \le 3,\ v \in V(G)} \mathsf{out}(H_{\mathrm{f}}(v, j)).$$

We construct a set $\mathcal{P}$ of pairs $\{z, z'\} \subseteq Z$ such that

(i) for every $\{z, z'\} \in \mathcal{P}$ there is some $j \in \{1, 2, 3\}$ and an edge $vw \in E(G)$ such that $z \in \mathsf{out}(H_{\mathrm{f}}(v, j))$ and $z' \in \mathsf{out}(H_{\mathrm{f}}(w, j))$,
(ii) every $z \in Z$ belongs to exactly one pair in $\mathcal{P}$.

Such $\mathcal{P}$ exists, since by construction each $H_{\mathrm{f}}(j, v)$ has $d_G(v)$ outputs; moreover, $\mathcal{P}$ can be obtained efficiently by a straight forward approach. Finally, we obtain a bipartite graph $H_G$ from the disjoint union of all $H(v)$, $v \in V(G)$, by adding for each pair $\{z, z'\} \in \mathcal{P}$ a new vertex $p_{z,z'} \in V_2(H_G)$ and the edges $zp_{z,z'}$ and $z'p_{z,z'}$.

**Lemma A.2** *For every graph $G$ with minimum degree 2 the following statements are equivalent.*

(i) *$G$ is 3-colorable;*
(ii) *$H_G$ has a total cover $\mathcal{C}$ such that elements of $\mathcal{C}$ are mutually disjoint, and $|X \cap V_1(H_G)| \le 2$ for all $X \in \mathcal{C}$;*
(iii) *$H_G$ has a total cover.*

**PROOF.** (i)$\Rightarrow$(ii). Let $\chi : V(G) \to \{1, 2, 3\}$ be a proper coloring of $G$. We define for each $v \in V(G)$ a cover $\mathcal{C}(v)$ of $H(v)$ as follows. Consider $v \in V(G)$ and put $j := \chi(v)$. For the color gadget $H_{\mathrm{c}}(v)$, we take the unique cover $\mathcal{C}_j$ which only exposes $\mathsf{out}_j(H_{\mathrm{c}}(v))$. To cover $\mathsf{out}_j(H_{\mathrm{c}}(v))$, we have no choice but using the set $X := \{\mathsf{out}_j(H_{\mathrm{c}}(v)), \mathsf{in}(H_{\mathrm{f}}(v, j))\}$. For the fan-out gadgets $H_{\mathrm{f}}(v, i)$, $i = 1, 2, 3$, we define total covers $\mathcal{C}^{(i)}$ according to Lemma A.1 as follows. For $i \ne j$ we choose $\mathcal{C}^{(i)}$ such that all outputs of $H_{\mathrm{f}}(v, i)$ are exposed; $\mathcal{C}^{(j)}$, however, must expose $\mathsf{in}(H_{\mathrm{f}}(v, j))$, since $\mathsf{in}(H_{\mathrm{f}}(v, j))$ is already contained in $X$. Therefore we must take for $\mathcal{C}^{(j)}$ the unique cover which covers all outputs of $H_\mathrm{f}(v,j)$ (cf. Lemma A.1). Thus we have defined a total cover

$$\mathcal{C}(v) := \mathcal{C}_j \cup \{X\} \cup \mathcal{C}^{(1)} \cup \mathcal{C}^{(2)} \cup \mathcal{C}^{(3)} \tag{A.1}$$

of $H(v)$ which exposes all outputs of $H_\mathrm{f}(v,i)$ for $i \neq j = \chi(v)$.

Consequently, $\bigcup_{v\in V(G)} \mathcal{C}(v)$ is a cover of $H_G$ which covers all vertices in

$$\bigcup_{v\in V(G)} V_2(H(v)) = V_2(H_G) \setminus \{\, p_{z,z'} \mid \{z,z'\} \in \mathcal{P} \,\}.$$

Now consider an arbitrary vertex $p_{z,z'}$ for $\{z,z'\} \in \mathcal{P}$; by definition, $z \in \mathsf{out}(H_\mathrm{f}(v,i))$ and $z' \in \mathsf{out}(H_\mathrm{f}(w,i))$ for some $i \in \{1,2,3\}$ and $vw \in E(G)$. Since $\chi$ is a proper coloring, $\chi(v) \neq \chi(w)$. Hence $i \neq \chi(v)$ or $i \neq \chi(w)$; w.l.o.g., we assume $i \neq \chi(v)$. It follows that $\mathcal{C}(v)$ exposes the outputs of $H_\mathrm{f}(v,i)$, and so we can use $X_{z,z'} := \{z, p_{z,z'}\}$ to cover $p_{z,z'}$. Whence

$$\mathcal{C} := \bigcup_{v\in V(G)} \mathcal{C}(v) \cup \{\, X_{z,z'} \mid \{z,z'\} \in \mathcal{P} \,\}$$

is a total cover of $H_G$. By construction, elements of $\mathcal{C}$ are mutually disjoint, and $|X \cap V_1(H_G)| \leq 2$ for all $X \in \mathcal{C}$.

(ii)⇒(iii). Trivial.

(iii)⇒(i). Let $\mathcal{C}$ be a total cover of $H_G$. For every $v \in V(G)$, the cover $\mathcal{C} \restriction H_\mathrm{c}(v)$ exposes at least one output of $H_\mathrm{c}(v)$. Hence we can define a coloring $\chi : V(G) \to \{1,2,3\}$ by choosing $\chi(v)$ such that $\mathsf{out}_{\chi(v)}(H_\mathrm{c}(v))$ is exposed by $\mathcal{C} \restriction H_\mathrm{c}(v)$. We will show that $\chi$ is a proper coloring. Suppose to the contrary that $\chi(v) = \chi(w) = j$ for some $vw \in E(G)$. Since $\mathsf{out}_j(H_\mathrm{c}(v))$ is exposed by $\mathcal{C} \restriction H_\mathrm{c}(v)$, it follows that $\{\mathsf{out}_j(H_\mathrm{c}(v)), \mathsf{in}(H_\mathrm{f}(v,j))\} \in \mathcal{C}$; consequently, by Lemma A.1, all outputs of $H_\mathrm{f}(v,j)$ are covered by $\mathcal{C} \restriction H_\mathrm{f}(v,j)$. The same argument shows that all outputs of $H_\mathrm{f}(w,j)$ are covered by $\mathcal{C} \restriction H_\mathrm{f}(w,j)$. By construction of $H_G$, there is a vertex $p_{z,z'} \in V_2(H_G)$ of degree 2 which is adjacent with some output $z$ of $H_\mathrm{f}(v,j)$ and some output $z'$ of $H_\mathrm{f}(w,j)$. Since $z$ and $z'$ are covered by $\mathcal{C} \restriction H_\mathrm{f}(v,j)$ and $\mathcal{C} \restriction H_\mathrm{f}(w,j)$, respectively, it follows that $p_{z,z'}$ is exposed by $\mathcal{C}$, a contradiction, as $\mathcal{C}$ is assumed to be a total cover. □

Clearly the problem whether a bipartite graph has a total cover is in NP. Moreover, the construction of $H_G$ can be carried out in polynomial time w.r.t. the size of $H$; hence Lemma A.2 implies the following.

**Theorem A.1** *Recognition of bipartite graphs $H$ having a total cover is* NP*-complete. This problem remains* NP*-complete if we only consider such covers $\mathcal{C}$ whose elements are mutually disjoint and $|X \cap V_2(H)| \leq 3$ for all $X \in \mathcal{C}$.*

Thus Theorems 1 and 2 are established.

## Appendix B. $\Pi_2^P$-completeness of graph satisfiability

This appendix is devoted to a proof of Theorem 3.

Let $G$ be a graph and $S \subseteq V(G)$. We say that $G$ is *$S$-extending 3-colorable* if every coloring $\chi_S : S \to \{1, 2\}$ can be extended to a proper 3-coloring of $G$. (If $S$ contains a pair of adjacent vertices, then $G$ is evidently not $S$-extending 3-colorable; hence it suffices to consider $S$ independent.)

**Theorem B.1** *The problem whether a graph $G$ with $S \subseteq V(G)$ is $S$-extending 3-colorable is $\Pi_2^P$-complete.*

**PROOF.** The problem belongs to $\Pi_2^P$, since if $G$ is not $S$-extending 3-colorability, then we can guess a 3-coloring $\chi_S$ of $S$ and check by a co-NP-oracle whether $\chi_S$ cannot be extended to a 3-coloring of $G$.

For $\Pi_2^P$-hardness we use a reduction from the following problem (NAE∀∃3SAT) which is shown $\Pi_2^P$-complete in [3] ("nae" stands for "not-all-equal").

> Given a quantified boolean formula $\Phi = \forall y_1, \dots, y_r \exists x_1, \dots, x_s F$, where $F$ is a CNF formula with exactly three literals in every clause; is $\Phi$ nae-satisfiable? (i.e., for every truth assignment to $y_1, \dots, y_r$ is there a truth assignment to $x_1, \dots, x_s$ such that every clause in $F$ contains both a true and a false literal?).

Consider $\Phi = \forall y_1, \dots, y_r \exists x_1, \dots, x_s F$, as above. We obtain from $F$ a graph $G$ by a standard transformation used elsewhere to reduce NAESAT to 3-colorability (see, e.g., [12, Theorem 9.8]). For each variable $z$ of $F$ we take two vertices $v_z$ and $v_{\overline{z}}$, and for each clause $C = (l_1 \vee l_2 \vee l_3)$ we take three vertices $v_{C,1}, v_{C,2}, v_{C,3}$; finally, we take one more vertex $a$. We add edges such that $v_z, v_{\overline{z}}, a$ forms a triangle for every variable $z$, and $v_{C,1}, v_{C,2}, v_{C,3}$ forms a triangle for every clause $C$. Finally, for every clause $C = (l_1 \vee l_2 \vee l_3)$ and $i = 1, 2, 3$, we join $v_{C,i}$ and $v_{l_i}$ by an edge. (See Figure B.1 for an illustration.)

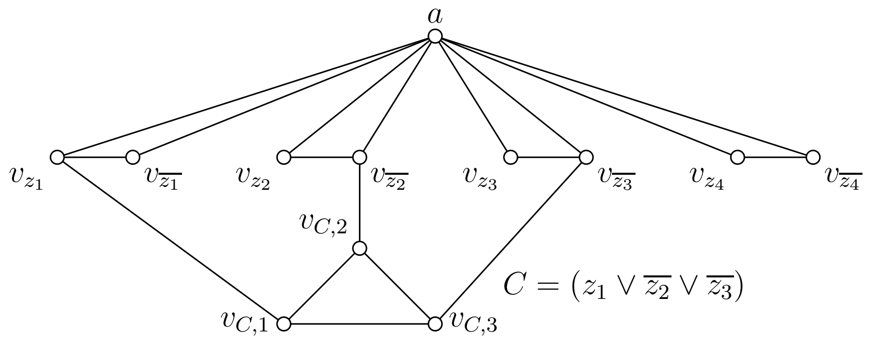


Fig. B.1. Reduction from nae-satisfiability.

$F$ is nae-satisfiable if and only if $G$ is 3-colorable (cf. [12, Theorem 9.8]). In fact, every nae-satisfying assignment of $F$ corresponds to a proper 3-coloring of $G$ such that vertex $v_l$ has color 1 if and only if $l$ is set TRUE by the assignment (fixing the color of vertex $a$ to 3.)

Now we put $S := \{v_{y_1}, \ldots, v_{y_r}\}$. We conclude that $\Phi$ is nae-satisfiable if and only if every coloring of $S$ by colors $\{1, 2\}$ can be extended to a proper 3-coloring of $G$. $\square$

Let $G$ be a graph with minimum degree 2 and $S \subseteq V(G)$. We construct a graph $H_{G,S}$ as in Appendix A, except for vertices in $S$ we use 2-*color gadgets* (defined below) instead of ordinary color gadgets (thus $H_{G,\emptyset}$ is nothing but $H_G$).

A 2-color gadget is a bipartite graph $H_{2c}$ with $V_1(H_{2c}) = \{x_1, x_2\}$ and $V_2(H_{2c}) = \{y_1, y_2, w_1, w_2\}$ (see Figure B.2). Similarly as for color gadgets, we call $\mathsf{out}_i(H_{2c}) = y_i$, $i = 1, 2$, the *outputs* of $H_{2c}$.

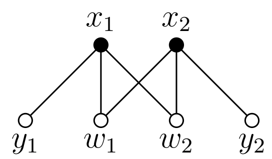


Fig. B.2. 2-color gadget $H_{2c}$.

In the following lemma we use the notation $x^1 := x$, $x^0 := \overline{x}$.

**Lemma B.1** *Let* $F = (x_1^{\alpha_1}) \wedge (x_1^{\alpha_2} \vee x_2^{\beta_3}) \wedge (x_1^{\alpha_3} \vee x_2^{\beta_2}) \wedge (x_2^{\beta_1})$, $\alpha_i, \beta_i \in \{0, 1\}$, *be a CNF formula (i.e.,* $G(F) = H_{2c}$*), and let* $F_1, F_2$ *be the result of removing the first or last clause from* $F$*, respectively. If* $F_1$ *and* $F_2$ *are both satisfiable, then* $F$ *is satisfiable.*

**PROOF.** If $\alpha_1 = \alpha_2 = \alpha_3$ or $\beta_1 = \beta_2 = \beta_3$, then $F$ is evidently satisfiable. The remaining cases are covered by nine combinations (a$i$,b$j$) according to

(a1) $\alpha_1 = \alpha_2 \neq \alpha_3$, (b1) $\beta_1 = \beta_2 \neq \beta_3$,

(a2) $\alpha_1 = \alpha_3 \neq \alpha_2$, (b2) $\beta_1 = \beta_3 \neq \beta_2$,

(a3) $\alpha_2 = \alpha_3 \neq \alpha_1$, (b3) $\beta_2 = \beta_3 \neq \beta_1$.

If $\alpha_2 \neq \alpha_3$ and $\beta_2 \neq \beta_3$, then $F$ is satisfiable, thus cases (a1,b1), (a1,b2), (a2,b1), and (a2,b2) are excluded. In case (a3,b3), $F$ is unsatisfiable, but $F_1$ and $F_2$ are unsatisfiable as well. Hence we are left with four cases. In cases (a1,b3) and (a2,b3), $F_1$ is unsatisfiable and $F_2$ is satisfiable. Symmetrically, in cases (a3,b1) and (a3,b2), $F_2$ is unsatisfiable and $F_1$ is satisfiable. $\square$

Let $G$ be a graph, $S \subseteq V(G)$, and $H_{G,S}$ the bipartite graph as defined above. Consider $v \in S$ and the 2-color gadget $H_{2c}(v)$. For $i \in \{1,2\}$ we write $f_{v,i}(H_{G,S})$ for the result of removing all vertices of $H_{2c}(v)$ from $H_{G,S}$ except the vertex $y_i = \mathsf{out}_i(H_{2c}(v))$; i.e., $f_{v,i}(H_{G,S}) := H_{G,S} - (V(H_{2c}(v)) \setminus \{y_i\})$. If $F$ is a CNF formula with $G(F) = H_{G,S}$, then we can define as well $F' = f_{v,i}(F)$ as the CNF formula which emerges from $F$ by the corresponding replacement of clauses; thus $G(F') = f_{i,v}(H_{G,S})$.

Moreover, we write $H^*_{G,S}$ for the result of applying this replacement for all $v \in S$, and call $H^*_{G,S}$ a *gadget reduction* of $H_{G,S}$. Note that $H_{G,S}$ has $2^{|S|}$ different gadget reductions.

The following is a direct consequence of Lemma B.1.

**Lemma B.2** *Let $F$ be a CNF formula with $G(F) = H_{G,S}$, and let $v \in S$. If both $f_{v,1}(F)$ and $f_{v,2}(F)$ are satisfiable, then $F$ is satisfiable.*

**Lemma B.3** *Let $G$ be a graph with minimum degree 2, $S \subseteq V(G)$ an independent set of vertices, and let $H_{G,S}$ be the corresponding bipartite graph as defined above. Then $G$ is $S$-extending 3-colorable if and only if every CNF formula $F$ with $G(F) = H_{G,S}$ is satisfiable.*

**PROOF.** ($\Rightarrow$). Assume that $G$ is $S$-extending 3-colorable, and choose any $F$ with $G(F) = H_{G,S}$. Assume to the contrary that $F$ is unsatisfiable. We conclude by Lemma B.2 that there is a gadget reduction $H^*_{G,S}$ of $H_{G,S}$ and some unsatisfiable CNF formula $F^*$ with $G(F^*) = H^*_{G,S}$.

We define a coloring $\chi_S$ of $S$ by setting $\chi_S(v) = i$ if and only if $f_{v,i}$ was used by obtaining $H^*_{G,S}$ from $H_{G,S}$. Since $G$ is $S$-extending 3-colorable, there is a proper 3-coloring $\chi$ of $G$ which extends $\chi_S$. We obtain a cover

$$\mathcal{C}' := \bigcup_{v \in V(G)} \mathcal{C}(v) \cup \{\, X_{z,z'} \mid \{z,z'\} \in \mathcal{P} \,\}$$

of $H_{G,S}$ similarly as in the first part of the proof of Lemma A.2; except for for $v \in S$ we put

$$\mathcal{C}(v) := \{X\} \cup \mathcal{C}^{(1)} \cup \mathcal{C}^{(2)},$$

where $X := \{\mathsf{out}_{\chi(v)}(H_{2c}(v)), \mathsf{in}(H_f(v,\chi(v)))\}$. In fact, $\mathcal{C}'$ is a total cover of $H^*_{G,S}$. Since $G(F^*) = H^*_{G,S}$, it follows that $F^*$ is graph satisfiable (and so satisfiable), a contradiction. Whence $F$ is indeed satisfiable.

($\Leftarrow$). Assume that every CNF formula $F$ with $G(F) = H_{G,S}$ is satisfiable, and choose an arbitrary coloring $\chi_S : S \to \{1,2\}$. We show that $\chi_S$ can be extended to a proper 3-coloring of $G$. We define a CNF formula $F$ with $G(F) = H_{G,S}$ as follows (naming of variables is chosen w.r.t. Figures A.1, A.2, and B.2).

- For every 2-color gadget $H_{2c}(v)$, $v \in S$, we take clauses

$$(\overline{x_1} \vee a^{(1)}), (x_1 \vee x_2), (x_1 \vee \overline{x_2}), (x_2 \vee a^{(2)}) \text{ if } \chi_S(v) = 1; \tag{B.1}$$
$$(x_1 \vee a^{(1)}), (\overline{x_1} \vee x_2), (x_1 \vee x_2), (\overline{x_2} \vee a^{(2)}) \text{ if } \chi_S(v) = 2 \tag{B.2}$$

  where $a^{(i)} = \mathsf{in}(H_f(v,i))$, $i = 1, 2$.
- For every color gadget $H_f(v)$ of $H_G$ we take clauses

$$(\overline{x_1} \vee \overline{x_2} \vee a^{(1)}), (\overline{x_1} \vee x_2 \vee a^{(2)}), (x_1 \vee \overline{x_2} \vee a^{(3)}), (x_1 \vee x_2), \tag{B.3}$$

  where $a^{(i)} = \mathsf{in}(H_f(v,i))$, $i = 1, 2, 3$.
- For every fan-out gadget $H_f(v,i)$ we take clauses of the form

$$\begin{gathered}(\overline{a_1} \vee b_1), (\overline{a_1} \vee \overline{b_1} \vee z_1), (\overline{a_1} \vee \overline{b_1} \vee a_2),\\ \cdots \\ (\overline{a_{q-1}} \vee b_{q-1}), (\overline{a_{q-1}} \vee \overline{b_{q-1}} \vee z_{q-1}), (\overline{a_{q-1}} \vee \overline{b_{q-1}} \vee z_q)\end{gathered} \tag{B.4}$$

  where $\mathsf{in}(H_f(v,i)) = a_1$, $\mathsf{out}(H_f(v,i)) = \{z_1, \ldots, z_q\}$.
- For every vertex $p_{z,z'} \in V(H_G)$ with $z \in \mathsf{out}(H_f(v,j))$ and $z' \in \mathsf{out}(H_f(w,j))$, we take the clause

$$(\overline{z} \vee \overline{z'}).$$

Since $F$ is assumed to be graph satisfiable, there is a truth assignment $t$ which satisfies $F$.

- Consider clauses in $F$ defined by (B.1). It follows that $t(x_1) = \text{TRUE}$, otherwise one of the clauses $(x_1 \vee x_2)$, $(x_1 \vee \overline{x_2})$, remains unsatisfied. Consequently $t(a^{(1)}) = \text{TRUE}$. A symmetric argument shows $t(a^{(2)}) = \text{TRUE}$ for clauses defined by (B.2).
- Consider clauses in $F$ defined by (B.3). Clearly $t(x_1) = \text{TRUE}$ or $t(x_2) = \text{TRUE}$ in order to satisfy $(x_1 \vee x_2)$; consequently $t(a^{(i)}) = \text{TRUE}$ for some $i \in \{1, 2, 3\}$.
- Consider clauses in $F$ defined by (B.4). If $t(a_1) = \text{TRUE}$ then necessarily $t(z_i) = \text{TRUE}$, $i = 1, \ldots, q$.

We extend $\chi_S$ to a 3-coloring $\chi : V(G) \to \{1, 2, 3\}$, choosing $\chi(v)$ for $v \in V(G) \setminus S$ as specified in the following table.

| $t(x_1)$ | $t(x_2)$ | $\chi(v)$ |
|---|---|---|
| TRUE | TRUE | 1 |
| TRUE | FALSE | 2 |
| FALSE | TRUE | 3 |

In view of the above observations, it follows that $t(z) = \text{TRUE}$ for all $z \in \mathsf{out}(H_{\mathrm{f}}(v, \chi(v)))$, $v \in V(G)$. It remains to show that $\chi$ is a proper 3-coloring. Suppose to the contrary that there is a pair of adjacent vertices $v, w \in V(G)$ with $\chi(v) = \chi(w) = j$. By construction of $H_{G,S}$, there is a vertex $p_{z,z'} \in V(H_G)$ of degree 2 which is incident with some $z \in \mathsf{out}(H_{\mathrm{f}}(v, j))$ and $z' \in \mathsf{out}(H_{\mathrm{f}}(w, j))$; thus $F$ contains the clause $(\overline{z} \vee \overline{z'})$. However, $t(z) = t(z') = \text{TRUE}$, since $\chi(v) = \chi(w) = j$. Hence $(\overline{z} \vee \overline{z'})$ is not satisfied, a contradiction. Since the coloring $\chi_S$ of $S$ was chosen arbitrarily, $G$ is indeed $S$-extending 3-colorable. $\square$

The combination of Lemma B.3 and Theorem B.1 yields $\Pi_2^{\mathrm{P}}$-hardness of graph satisfiability. Thus Theorem 3 is shown true.