ELSEVIER

Artificial Intelligence 107 (1999) 219–263

Artificial Intelligence

Logical analysis of binary data with missing bits ☆

Endre Boros [a,1], Toshihide Ibaraki [b,*], Kazuhisa Makino [c,2]

[a] *RUTCOR, Rutgers University, 640 Bartholomew Road, Piscataway, NJ 08854-8003, USA*
[b] *Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan*
[c] *Department of Systems and Human Science, Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka 560-8531, Japan*

Received 21 November 1997

Abstract

We model a given pair of sets of positive and negative examples, each of which may contain missing components, as a partially defined Boolean function with missing bits (pBmb) $(\widetilde{T}, \widetilde{F})$, where $\widetilde{T} \subseteq \{0, 1, *\}^n$ and $\widetilde{F} \subseteq \{0, 1, *\}^n$, and "$*$" stands for a missing bit. Then we consider the problem of establishing a Boolean function (an extension) $f : \{0, 1\}^n \to \{0, 1\}$ belonging to a given function class $\mathcal{C}$, such that f is true (respectively, false) for every vector in $\widetilde{T}$ (respectively, in $\widetilde{F}$). This is a fundamental problem, encountered in many areas such as learning theory, pattern recognition, example-based knowledge bases, logical analysis of data, knowledge discovery and data mining. In this paper, depending upon how to deal with missing bits, we formulate three types of extensions called robust, consistent and most robust extensions, for various classes of Boolean functions such as general, positive, Horn, threshold, decomposable and k-DNF. The complexity of the associated problems are then clarified; some of them are solvable in polynomial time while the others are NP-hard. © 1999 Elsevier Science B.V. All rights reserved.

Keywords: Knowledge discovery; Data mining; Logical analysis of data; Boolean functions; Partially defined Boolean functions; Missing bits; NP-hardness

☆ This research was partially supported by ONR (Grants N00014-92-J-1375 and N00014-92-J-4083), and the Scientific Grants in Aid by the Ministry of Education, Science, Sports and Culture of Japan. The visit of the first author to Kyoto University was made possible by the grant (06044112) of the Ministry of Education, Science, Sports and Culture of Japan.

* Corresponding author. Email: ibaraki@kuamp.kyoto-u.ac.jp.

[1] Email: boros@rutcor.rutgers.edu.

[2] Email: makino@sys.es.osaka-u.ac.jp.

PII: S0004-3702(98)00110-6

1. Introduction

In analyzing data of some phenomena from a logical viewpoint, we often encounter the following problem: Given a pair of data sets (T, F) of "positive" and "negative" examples, where $T, F \subseteq \{0, 1\}^n$, establish a Boolean function (extension) f in a specified function class $\mathcal{C}$, such that f is true (respectively, false) for every vector in T (respectively, in F). A pair of sets (T, F) is called a *partially defined Boolean function (pdBf)*.

For instance, a data vector x may represent the symptoms to diagnose a disease, e.g., x_1 denotes whether temperature is high ($x_1 = 1$) or not ($x_1 = 0$), and x_2 denotes whether blood pressure is high ($x_2 = 1$) or not ($x_2 = 0$), etc. Establishing an *extension* f, which is consistent with the given data set, then amounts to finding its logical diagnostic explanation.

This type of problems is studied, for example, in learning theory (e.g., [3,32,38]), where it is called the *consistency problem*. In the process of learning, it is fundamental to find an extension of the current set of data (T, F). The learner tries to find an extension of small (i.e., polynomial) size as it leads to interesting theoretical consequences [7]. In pattern recognition, a function separating two categories of data T and F is usually called a discriminant function (e.g., [28]). If the data are binary, this is essentially the same as an extension of a pdBf (T, F). In example-based knowledge bases, we encounter a similar problem of establishing an extension, but in this case it is usually asked to describe the extension by rules. Finding extensions is also one of the main goals in such areas as data analysis, knowledge acquisition, knowledge discovery and data mining (e.g., [1,9,16,17, 34]), which are recently receiving increasing attention.

In many of the above applications, some knowledge or hypothesis about the extension f is usually available beforehand. Such knowledge may be obtained from experience or from the analysis of mechanisms that may or may not cause the phenomena under consideration. In the above example of diagnosing diseases, it would be natural to assume that we somehow know the direction of each variable that tends to cause the disease to appear. By changing the polarities of variables if necessary, therefore, the extension $f(x)$ can be assumed to be positive (i.e., monotone increasing) in all variables.

As the above observation is essential, we consider in this paper to find an extension f that belongs to a specified class of functions $\mathcal{C}$. The classes of functions considered in this paper include general, positive (or monotone), Horn, threshold, decomposable and k-DNF. The class of positive functions may be the most natural special class to investigate in this respect. Horn functions are important in the sense that the satisfiability problem of Horn CNF (conjunctive normal form) can be solved in polynomial time [4,19], and, for this reason, logic programs and expert systems are often built on Horn rules. If an extension f is Horn, its true set (or false set, depending on the definition) can be described by a Horn CNF. Threshold functions [30] have an appealing geometrical interpretation of linear separation, and hence is a major tool to describe discriminant functions used in pattern recognition (e.g., [28]). Decomposable functions [5,8,35] are important because they can provide us additional information regarding the hierarchical structure underlying the given data sets. Finally the class of k-DNF should also be included in the list, since DNF is a standard form of representation of Boolean functions. The prime implicants in DNF of an extension are also called "association rules" in data mining (e.g., [1,29]), and "patterns" in papers on logical analysis of data [16] and its applications [10].

Unfortunately, real-world data might not be complete, adding another dimension of complication. In other words, the values of some elements x_j in a given data vector x may not be available for various reasons, such as the test to measure the x_j was not conducted because it takes too much time or is expensive, or the data bits are simply lost. Therefore, it is indispensable to admit incomplete data in order to be usable in practical applications. We denote the missing bits by "$*$" in this paper. A set of data $(\widetilde{T}, \widetilde{F})$, which includes missing bits, is called a *partially defined Boolean function with missing bits (pBmb)*, where $\widetilde{T} \subseteq \{0, 1, *\}^n$ (respectively, $\widetilde{F} \subseteq \{0, 1, *\}^n$) denotes the set of "positive examples" (respectively, "negative examples").

We introduce in this paper three types of extensions of a pBmb $(\widetilde{T}, \widetilde{F})$, called robust, consistent and most robust extensions, depending upon how we deal with the missing bits. More precisely, given a pBmb $(\widetilde{T}, \widetilde{F})$, a Boolean function f is called (i) a *robust* extension if for every $\tilde{a} \in \widetilde{T}$ (respectively, $\tilde{a} \in \widetilde{F}$), any 0–1 vector a obtained from $\tilde{a}$ by fixing its missing bits arbitrarily satisfies $f(a) = 1$ (respectively, $f(a) = 0$). It is called (ii) a *consistent* extension if for every $\tilde{a} \in \widetilde{T}$ (respectively, $\tilde{a} \in \widetilde{F}$), there exists a 0–1 vector a obtained from $\tilde{a}$ by fixing its missing bits appropriately, for which $f(a) = 1$ (respectively, $f(a) = 0$) holds. Finally, f is called (iii) a *most robust* extension if it is a robust extension of the pBmb (T', F') obtained from $(\widetilde{T}, \widetilde{F})$ by fixing a smallest set of missing bits appropriately (the remaining missing bits in $T' \cup F'$ are assumed to take arbitrary values).

All of these extensions provide logical explanations of a given pBmb $(\widetilde{T}, \widetilde{F})$ with varied freedom given to the missing bits in $\widetilde{T}$ and $\widetilde{F}$. Let us remark that by definition, if $f \in \mathcal{C}$ is a most robust extension of $(\widetilde{T}, \widetilde{F})$, then it is also consistent; furthermore, if $(\widetilde{T}, \widetilde{F})$ has a robust extension in a class $\mathcal{C}$, then it also has a most robust one (and hence a consistent one, too). Let us add that the process of finding consistent and most robust extensions will also provide us with conditions on the values of missing bits, required for $(\widetilde{T}, \widetilde{F})$ to have a consistent extension in the given class $\mathcal{C}$. This type of information can also be useful in analyzing incomplete data sets.

In the above example of diagnosing diseases, not all medical tests are usually performed on each patient, because the tests may be painful, expensive or even dangerous. Such attributes thus naturally become missing. In this case, a robust extension provides very useful information, if it exists, since it is a diagnostic explanation of the disease under consideration regardless of the interpretation of the missing bits. That is, it says that the current data set carries enough information to derive a meaningful explanation. It may happen, however, that the data set has no robust extension. Even in this case, there may be an extension if we can supply correct interpretation of all or part of the missing bits; such an extension may help or even improve diagnostic procedures. This leads to the concepts of consistent and most robust extensions. The most robust extension is important in practice as it minimizes the number of "corrected" bits in order to have an extension. There are other possible treatments of missing bits appearing in the context of learning theory (see, e.g., [6,18,23,25,36,37,39]).

In this paper, we study the problems of deciding the existence of these extensions for various special classes of Boolean functions $\mathcal{C}$, mainly from the viewpoint of their computational complexity. We obtain computationally efficient algorithms in some cases, and prove NP-hardness in some other cases. For all the cases with efficient algorithms, we also provide efficient algorithms to construct the corresponding extensions. It should also be interesting to note that all such extensions are short (i.e., of linear size of input length). For a summary of the results obtained, see Tables 1 and 2 in the last section. A part of these results was already presented in [13,14], and a more comprehensive discussion can be found in the technical report [12].

2. Preliminaries

2.1. Boolean functions

A *Boolean function*, or a *function* in short, is a mapping $f:\mathbb{B}^n \mapsto \mathbb{B}$, where $\mathbb{B}=\{0,1\}$, and $x \in \mathbb{B}^n$ is called a *Boolean vector* (or a *vector* in short). If $f(x)=1$ (respectively, 0), then x is a *true* (respectively, *false*) vector of f. The set of all true (respectively, false) vectors is denoted by $T(f)$ (respectively, $F(f)$). Two special functions with $T(f)=\emptyset$ and $F(f)=\emptyset$ are, respectively, denoted by $f=\bot$ and $f=\top$. For two functions f and g on the same set of variables, we write $f \leqslant g$ if $f(x)=1$ implies $g(x)=1$ for all $x \in \mathbb{B}^n$, and we write $f<g$ if $f \leqslant g$ and $f \neq g$.

A function f is called *positive* if $x \leqslant y$ (i.e., $x_i \leqslant y_i$ for all $i \in \{1,2,\ldots,n\}$) always implies $f(x) \leqslant f(y)$. A positive function is also called *monotone*. The variables $x_1, x_2, \ldots, x_n$ and their complements $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n$ are called *literals*.

A *term* is a conjunction of literals such that at most one of x_i and $\bar{x}_i$ appears for each i. The constant 1 (viewed as the conjunction of an empty set of literals) is also considered as a term. A *disjunctive normal form* (DNF) is a disjunction of terms. Clearly, a DNF defines a function, and it is well known that every function can be represented by a DNF (however, such a representation may not be unique). Throughout this paper, unless otherwise stated, we usually do not distinguish a DNF φ from the function it represents.

It is well known that a Boolean function f is positive if and only if f can be represented by a DNF, in which all the literals in each of the terms are uncomplemented. We shall call f a k-DNF if it has a DNF with at most k literals in each term, and it will be called *Horn* if it has a DNF with at most one complemented literal in each term. Let us denote, respectively, by $\mathcal{C}_{\mathrm{all}}$, $\mathcal{C}^+$, $\mathcal{C}_{k\text{-DNF}}$, $\mathcal{C}^+_{k\text{-DNF}}$ and $\mathcal{C}_{\mathrm{Horn}}$ the classes of all, positive, k-DNF, positive k-DNF and Horn Boolean functions. Note that, when we say k-DNF, k is considered to be general (i.e., it is a parameter included in the problem specification). However, since the problems may become easier to solve if k is fixed, we shall also investigate the cases of fixed k extensively.

A function f for which there exist $n+1$ real numbers $w_1, w_2, \ldots, w_n$ and t such that:

$$f(x)=\begin{cases} 1, & \text{if } \sum w_i x_i \geqslant t, \\ 0, & \text{if } \sum w_i x_i < t \end{cases} \tag{1}$$

is called *threshold*. Let us denote by $\mathcal{C}_{\mathrm{TH}}$ the family of threshold Boolean functions. It is known (see, e.g., [30]) that a threshold function f can equivalently be defined as

$$f(x)=\begin{cases} 1, & \text{if } \sum w_i' x_i \geqslant t', \\ 0, & \text{if } \sum w_i' x_i \leqslant t'-1 \end{cases} \tag{2}$$

for some (other) real numbers $w'_1, w'_2, \ldots, w'_n$ and t'. In this paper, we shall employ definition (2) instead of (1), in order to simplify the presentation of some of the proofs.

Let $V = \{1, 2, \ldots, n\}$ denote the index set of variables. For a vector $x \in \mathbb{B}^n$ and $S \subseteq V$, $x[S]$ denotes the projection of x on S. To simplify notation, for a Boolean function h depending only on variables of $S \subseteq V$, we write $h(S)$ instead of $h(x[S])$. For given $S_i \subseteq V$, $i = 0, 1, \ldots, k$, a function f is called $G(S_0, G(S_1), G(S_2), \ldots, G(S_k))$*-decomposable* (see, e.g., [5,8,27,35]), where G stands for the general Boolean functions, if there exist Boolean functions $h_1, \ldots, h_k$ and g satisfying the following three conditions:

(i) h_i depends only on variables in S_i, $i = 1, \ldots, k$,

(ii) g depends on the variables in S_0 and on the binary values $h_i(S_i)$ for $i = 1, \ldots, k$, (i.e., $g : \{0, 1\}^{|S_0|+k} \to \{0, 1\}$),

(iii) $f = g(S_0, h_1(S_1), h_2(S_2), \ldots, h_k(S_k))$.

Let us note that $S_0, S_1, \ldots, S_k$ are not necessarily assumed to be disjoint. Also, for given $S_i \subseteq V$, $i = 0, 1, \ldots, k$, a function f is called *positive* $G(S_0, G(S_1), G(S_2), \ldots, G(S_k))$*-decomposable* if f is $G(S_0, G(S_1), G(S_2), \ldots, G(S_k))$-decomposable, and the functions g and h_i, $i = 1, 2, \ldots, k$, are all positive. Let us denote by $\mathcal{C}_{G(S_0, G(S_1))}$ and $\mathcal{C}^+_{G(S_0, G(S_1))}$ the families of $G(S_0, G(S_1))$-decomposable and positive $G(S_0, G(S_1))$-decomposable functions, respectively.

2.2. *Partially defined Boolean functions and their extensions*

A *partially defined Boolean function (pdBf)* is defined by a pair of sets (T, F) such that $T, F \subseteq \mathbb{B}^n$. A function f is called an *extension* of the pdBf (T, F) if $T \subseteq T(f)$ and $F \subseteq F(f)$. We shall also say in this case that the function f *correctly classifies* all the vectors $a \in T$ and $b \in F$. Evidently, the disjointness of the sets T and F is a necessary and sufficient condition for the existence of an extension. It may not be evident, however, how to find out whether a given pdBf has a extension that belongs to a class of functions $\mathcal{C}$. Therefore, we have extensively studied the following problems in [11]. (As noted in Section 1, the first problem is called the consistency problem in learning theory, and some results obtained therein (e.g., [32]) overlap with those in [11].)

Problem EXTENSION($\mathcal{C}$)
Input: a pdBf (T, F), where $T, F \subseteq \mathbb{B}^n$.
Question: Is there an extension $f \in \mathcal{C}$ of (T, F)?

Problem BEST-FIT($\mathcal{C}$)
Input: a pdBf (T, F), where $T, F \subseteq \mathbb{B}^n$, and a weight function $w : T \cup F \mapsto \mathbb{R}_+$ (nonnegative reals).
Output: Subsets T^* and F^* such that $T^* \cap F^* = \emptyset$ and $T^* \cup F^* = T \cup F$, for which the pdBf (T^*, F^*) has an extension in $\mathcal{C}$, and $w(T \cap F^*) + w(F \cap T^*)$ is minimum.

We denote the minimum weight sum of erroneously classified vectors by a best-fit extension of (T, F) as

$$\varepsilon(T, F) = \min\{w(T \cap F^*) + w(F \cap T^*)\}. \tag{3}$$

As a pdBf does not allow missing bits, we shall introduce the set

$$\mathbb{M} = \{0, 1, *\},$$

and interpret the asterisk components $*$ of $v \in \mathbb{M}^n$ as missing bits. For a vector $v \in \mathbb{M}^n$, let $ON(v) = \{j \mid v_j = 1,\ j = 1, 2, \ldots, n\}$ and $OFF(v) = \{j \mid v_j = 0,\ j = 1, 2, \ldots, n\}$. For a subset $\widetilde{S} \subseteq \mathbb{M}^n$, let $AS(\widetilde{S}) = \{(v, j) \mid v \in \widetilde{S},\ j \in V \setminus (ON(v) \cup OFF(v))\}$ be the collection of all missing bits of the vectors in $\widetilde{S}$. If $\widetilde{S}$ is a singleton $\{v\}$, we also denote $AS(\{v\})$ as $AS(v)$. Clearly, $\mathbb{B}^n \subseteq \mathbb{M}^n$, and $v \in \mathbb{B}^n$ holds if and only if $AS(v) = \emptyset$. Let us consider a binary assignment $\alpha \in \mathbb{B}^Q$ to a subset $Q \subseteq AS(\widetilde{S})$ of the missing bits. Then v^α denotes the vector obtained from v by replacing the $*$ components which belong to Q by the binary values assigned by α, i.e.,

$$v_j^\alpha = \begin{cases} v_j & \text{if } (v, j) \notin Q, \\ \alpha(v, j) & \text{if } (v, j) \in Q, \end{cases}$$

and $\widetilde{S}^\alpha$ denotes $\{v^\alpha \mid v \in \widetilde{S}\}$. For example, $\widetilde{S} = \{u = (1, *, 0, 1), v = (0, 1, *, *), w = (1, 1, *, 0)\}$ has $AS(\widetilde{S}) = \{(u, 2), (v, 3), (v, 4), (w, 3)\}$. If $Q = \{(u, 2), (v, 4)\}$, an assignment $(\alpha(u, 2), \alpha(v, 4)) = (1, 0)$ yields $\widetilde{S}^\alpha = \{u^\alpha = (1, 1, 0, 1), v^\alpha = (0, 1, *, 0), w^\alpha = (1, 1, *, 0)\}$.

For vectors $v, w \in \mathbb{M}^n$, we shall write $v \gtrapprox w$ (respectively, $v \lessapprox w$) if there exists an assignment $\alpha \in \mathbb{B}^{AS(\{v, w\})}$ for which $v^\alpha \geqslant w^\alpha$ (respectively, $v^\alpha \leqslant w^\alpha$) holds, and we say that v is *potentially greater* (respectively, *smaller*) than w. If both $v \gtrapprox w$ and $v \lessapprox w$ hold then we write $v \approx w$, and say that v is *potentially identical* with w. Note that $v \approx w$ holds if and only if there is an assignment $\alpha \in AS(\{v, w\})$ such that $v^\alpha = w^\alpha$. For example, $v = (1, 0, *, 1, *)$ and $w = (0, 0, 0, 1, *)$ satisfy $v \gtrapprox w$, and $v' = (*, 0, 1, *, 1)$ and $w' = (1, 0, *, *, 1)$ satisfy $v' \approx w'$.

A *pdBf with missing bits* (or in short *pBmb*) is a pair $(\widetilde{T}, \widetilde{F})$, where $\widetilde{T}, \widetilde{F} \subseteq \mathbb{M}^n$. To a pBmb $(\widetilde{T}, \widetilde{F})$ we always associate the set $AS = AS(\widetilde{T} \cup \widetilde{F})$ of its missing bits. For a pBmb $(\widetilde{T}, \widetilde{F})$ and an assignment $\alpha \in \mathbb{B}^{AS}$, let $(\widetilde{T}^\alpha, \widetilde{F}^\alpha)$ be the pdBf defined by $\widetilde{T}^\alpha = \{a^\alpha \mid a \in \widetilde{T}\}$ and $\widetilde{F}^\alpha = \{b^\alpha \mid b \in \widetilde{F}\}$.

Let us call a function f a *robust extension* of the pBmb $(\widetilde{T}, \widetilde{F})$ if

$$f(a^\alpha) = 1 \text{ and } f(b^\alpha) = 0 \text{ for all } a \in \widetilde{T},\ b \in \widetilde{F} \text{ and for all } \alpha \in \mathbb{B}^{AS}.$$

We first consider the problem of deciding the existence of a robust extension of a given pBmb $(\widetilde{T}, \widetilde{F})$ in a specified class $\mathcal{C}$.

> Problem RE($\mathcal{C}$)
> Input: A pBmb $(\widetilde{T}, \widetilde{F})$, where $\widetilde{T}, \widetilde{F} \subseteq \mathbb{M}^n$.
> Question: Does $(\widetilde{T}, \widetilde{F})$ have a robust extension in class $\mathcal{C}$?

In case of YES, a robust extension $f \in \mathcal{C}$ must also be provided, either by a direct algebraic form, or by a polynomial time membership oracle. (A *membership oracle* for a function f is an algorithm that returns the value $f(v)$ for any given vector $v \in \mathbb{B}^n$.) Note that a vector $a \in \mathbb{M}^n$ can be seen as a subhypercube $\{a^\alpha \mid \alpha \in AS(a)\}$ of $\mathbb{B}^n$. Therefore, a robust extension can be regarded as an extension of two sets of hypercubes $\widetilde{T}$ and $\widetilde{F}$.

It may happen that a pBmb $(\widetilde{T}, \widetilde{F})$ has no robust extension in $\mathcal{C}$, but it has an extension if we change some (or all) the $*$ bits to appropriate binary values. A function f is called a *consistent extension* of pBmb $(\widetilde{T}, \widetilde{F})$, if there exists an assignment $\alpha \in \mathbb{B}^{AS}$ such that $f(a^\alpha) = 1$ and $f(b^\alpha) = 0$ for all $a \in \widetilde{T}$ and $b \in \widetilde{F}$. In other words, a pBmb $(\widetilde{T}, \widetilde{F})$ is said to have a consistent extension in $\mathcal{C}$ if, for some assignment $\alpha \in \mathbb{B}^{AS}$, the pdBf $(\widetilde{T}^\alpha, \widetilde{F}^\alpha)$ has an extension in $\mathcal{C}$.

Problem CE($\mathcal{C}$)
Input: A pBmb $(\widetilde{T}, \widetilde{F})$, where $\widetilde{T}, \widetilde{F} \subseteq \mathbb{M}^n$.
Question: Does $(\widetilde{T}, \widetilde{F})$ have a consistent extension in class $\mathcal{C}$?

In case of YES, an assignment $\alpha \in \mathbb{B}^{AS}$, for which the pdBf $(\widetilde{T}^\alpha, \widetilde{F}^\alpha)$ has an extension in $\mathcal{C}$, together with such an extension $f \in \mathcal{C}$, must also be provided.

Let us finally consider the case in which there is a consistent extension of the considered pBmb in the specified class, but for which not all missing bits should necessarily be specified. Let us call an assignment $\alpha \in \mathbb{B}^Q$ for a subset $Q \subseteq AS$ a *robust assignment* (with respect to a class $\mathcal{C}$) if the resulting pBmb $(\widetilde{T}^\alpha, \widetilde{F}^\alpha)$ has a robust extension in class $\mathcal{C}$. We are interested in finding such a robust assignment with the smallest size $|Q|$. Such an extension will be called a *most robust extension* of the given pBmb $(\widetilde{T}, \widetilde{F})$ in the class $\mathcal{C}$.

Problem MRE($\mathcal{C}$)
Input: A pBmb $(\widetilde{T}, \widetilde{F})$, where $\widetilde{T}, \widetilde{F} \subseteq \mathbb{M}^n$.
Output: NO if $(\widetilde{T}, \widetilde{F})$ does not have a consistent extension in class $\mathcal{C}$; otherwise a robust assignment $\alpha \in \mathbb{B}^Q$ for a subset $Q \subseteq AS$, which minimizes $|Q|$.

Similarly to the previous problems, if $(\widetilde{T}, \widetilde{F})$ has a consistent extension in class $\mathcal{C}$, a most robust extension $f \in \mathcal{C}$ of the pBmb $(\widetilde{T}^\alpha, \widetilde{F}^\alpha)$ must also be provided.

Let us denote the minimum size of $Q \subseteq AS$ having a robust assignment by

$$\rho\big(\mathcal{C}; (\widetilde{T}, \widetilde{F})\big) = \min_{\substack{Q \subseteq AS \\ \exists \alpha \in \mathbb{B}^Q \text{ s.t. } (\widetilde{T}^\alpha, \widetilde{F}^\alpha) \\ \text{has a robust extension in } \mathcal{C}}} |Q|, \tag{4}$$

where $\rho\big(\mathcal{C}; (\widetilde{T}, \widetilde{F})\big) = +\infty$ if there is no Q satisfying the stated condition. To simplify notation, we shall use sometimes $\rho(\widetilde{T}, \widetilde{F})$ in place of $\rho(\mathcal{C}; (\widetilde{T}, \widetilde{F}))$, unless confusion arises. Observe that a pBmb $(\widetilde{T}, \widetilde{F})$ has a robust extension if and only if $\rho(\widetilde{T}, \widetilde{F}) = 0$, and it has a consistent extension if and only if $\rho(\widetilde{T}, \widetilde{F}) \leqslant |AS|$.

It follows therefore that if RE($\mathcal{C}$) or CE($\mathcal{C}$) are NP-hard, then MRE($\mathcal{C}$) is NP-hard, and conversely, if MRE($\mathcal{C}$) is solvable in polynomial time, then both RE($\mathcal{C}$) and CE($\mathcal{C}$) are polynomially solvable. It seems also that RE($\mathcal{C}$) is, in general, easier than CE($\mathcal{C}$), since RE($\mathcal{C}$) can be seen as the extension problem of sets of hypercubes $\widetilde{T}$ and $\widetilde{F}$. Let us also note that, if $AS = \emptyset$ (i.e., $(\widetilde{T}, \widetilde{F})$ is a pdBf), then the notions of extension, robust extension and consistent extension all coincide. Thus, RE($\mathcal{C}$) and CE($\mathcal{C}$) are both at least as difficult as EXTENSION($\mathcal{C}$).

As we shall see in this paper that many of the above problems for various classes are NP-hard, we also extensively consider the following case:

$$|AS(a)| \leqslant k \quad \text{for all } a \in \widetilde{T} \cup \widetilde{F}, \tag{5}$$

where k is a positive constant. This is important because such constraints may often be met in real situations if the number of missing bits in the data is relatively small, and the problems then tend to become easier. For example, many of the data sets which appear in the Machine Learning Repository of the Computer Science Department of the University of California at Irvine [31] satisfy the above condition with a small constant k. Note that, in this case, the complexity of RE($\mathcal{C}$) is polynomially equivalent to that of EXTENSION($\mathcal{C}$). This is because a pBmb $(\widetilde{T}, \widetilde{F})$ has a robust extension if and only if the pdBf (T', F') has an extension in $\mathcal{C}$, where T' (respectively, F') is obtained by expanding each $a \in \widetilde{T}$ (respectively, $\widetilde{F}$) by assigning all possible $\alpha \in \mathbb{B}^{AS(a)}$ to the missing bits in a, and the size of (T', F') is at most 2^k-times the size of $(\widetilde{T}, \widetilde{F})$. Furthermore, we shall show that, for several classes $\mathcal{C}$, if $k = 1$ in (5), then the problem is tractable, but becomes intractable (unless P=NP) if $k \geqslant 2$ holds; i.e., $k = 1$ is a critical value of the problem. In fact, all the problems considered in this paper has either no critical value or a critical value of $k = 1$.

In this paper, we consider all function classes $\mathcal{C}$ defined in Section 2.1. Further interesting classes, such as regular, unate, renamable Horn, dual-minor, dual-major, self-dual, read-once and h-term k-DNFs are discussed in [12].

3. Problems RE and CE

In this section, we study the decision problems CE($\mathcal{C}$) and RE($\mathcal{C}$) for various classes of functions $\mathcal{C}$. Let us point out first that a basic difference between these two problems is in the verification of a positive answer. On one hand, it is easy to see that problem CE($\mathcal{C}$) belongs to NP, whenever problem EXTENSION($\mathcal{C}$) is in NP. Namely, a pBmb $(\widetilde{T}, \widetilde{F})$ has a consistent extension if the pdBf $(\widetilde{T}^{\alpha}, \widetilde{F}^{\alpha})$ has an extension in $\mathcal{C}$ for only one assignment $\alpha \in \mathbb{B}^{AS}$, and this can be accomplished in nondeterministic polynomial time by the assumption that EXTENSION($\mathcal{C}$) belongs to NP. On the other hand, problem RE($\mathcal{C}$) may not belong to NP, due to the condition "for all $\alpha \in \mathbb{B}^{AS}$" which a robust extension f must satisfy. For example, if $f \in \mathcal{C}$ is a robust extension of a pBmb $(\widetilde{T}, \widetilde{F})$, and $a \in \widetilde{T}$, then checking the equation $f(a^{\alpha}) = 1$ for all $\alpha \in \mathbb{B}^{AS(a)}$ may amount to a tautology problem in those variables x_j with $(a, j) \in AS(a)$. In fact, we shall see later that RE($\mathcal{C}$) is co-NP-complete for some classes $\mathcal{C}$.

After summarizing implications of EXTENSION($\mathcal{C}$) in the next subsection, we consider problems RE($\mathcal{C}$) and CE($\mathcal{C}$) for respective classes $\mathcal{C}$ in the subsequent subsections.

3.1. *Implications by problem* EXTENSION

Let us observe that, as mentioned earlier, EXTENSION($\mathcal{C}$) is a special case of both RE($\mathcal{C}$) and CE($\mathcal{C}$). Hence we have the following theorem.

Theorem 1. *If problem* EXTENSION($\mathcal{C}$) *is* NP-*complete, then problem* CE($\mathcal{C}$) *is* NP-*complete, and problem* RE($\mathcal{C}$) *is* NP-*hard.*

The slight difference between the conclusions for CE($\mathcal{C}$) and RE($\mathcal{C}$) comes from the fact that problem RE($\mathcal{C}$) may not belong to class NP, as we pointed out it earlier. We immediately have the following corollary from the results of [11].

Corollary 1. *Problem* CE($\mathcal{C}$) *is* NP*-complete and problem* RE($\mathcal{C}$) *is* NP*-hard for the classes of k-DNF functions and positive k-DNF functions.*

We also can derive the following positive result.

Theorem 2. *If problem* EXTENSION($\mathcal{C}$) *can be solved in polynomial time, then problem* RE($\mathcal{C}$) *is also polynomially solvable for pBmb instances* $(\widetilde{T}, \widetilde{F})$ *satisfying* $|AS(a)| = \mathrm{O}(\log n)$ *for all* $a \in \widetilde{T} \cup \widetilde{F}$*, where* $\widetilde{T}, \widetilde{F} \subseteq \mathbb{M}^n$.

Proof. It follows from the definition that a pBmb $(\widetilde{T}, \widetilde{F})$ has a robust extension in $\mathcal{C}$ if and only if the pdBf (T', F') has an extension in $\mathcal{C}$, where T' and F' are defined by

$$T' = \{a^\alpha \mid a \in \widetilde{T},\ \alpha \in \mathbb{B}^{AS(a)}\}, \tag{6}$$

$$F' = \{b^\alpha \mid b \in \widetilde{F},\ \alpha \in \mathbb{B}^{AS(b)}\}. \tag{7}$$

Since $|T'| + |F'| = \mathrm{O}(n(|\widetilde{T}| + |\widetilde{F}|))$ holds by $|AS(a)| = \mathrm{O}(\log n)$ for all $a \in \widetilde{T} \cup \widetilde{F}$, the polynomiality of EXTENSION($\mathcal{C}$) then implies the polynomiality of RE($\mathcal{C}$). □

Corollary 2. *Let a pBmb* $(\widetilde{T}, \widetilde{F})$ *satisfy* $|AS(a)| = \mathrm{O}(\log n)$ *for all* $a \in \widetilde{T} \cup \widetilde{F}$*, where* $\widetilde{T}, \widetilde{F} \subseteq \mathbb{M}^n$*. For such instances problem* RE($\mathcal{C}$) *can be solved in polynomial time, if* $\mathcal{C}$ *is one of the following classes*: (1) *general,* (2) *positive,* (3) *Horn,* (4) *threshold,* (5) (*positive*) $g(S_0, h_1(S_1))$*-decomposable, and* (6) (*positive*) k-DNF *with fixed* k.

Proof. Combine Theorem 2 and the results in [11]. □

3.2. General extensions

Let us consider problems RE and CE for the class $\mathcal{C}_{\mathrm{all}}$ of all Boolean functions. We shall start with an easy result.

Theorem 3. *Problem* RE($\mathcal{C}_{\mathrm{all}}$) *can be solved in polynomial time.*

Proof. It is easy to see that a pBmb $(\widetilde{T}, \widetilde{F})$ has a robust extension if and only if for each $\alpha \in AS(\widetilde{T})$ and for each $\beta \in AS(\widetilde{F})$ the sets of binary vectors $\widetilde{T}^\alpha$ and $\widetilde{F}^\beta$ are disjoint, in other words, if and only if for every pair of $a \in \widetilde{T}$ and $b \in \widetilde{F}$ there exists an index j such that $a_j \neq b_j$ and $\{a_j, b_j\} = \{0, 1\}$ (i.e., either $a_j = 0$ and $b_j = 1$, or $a_j = 1$ and $b_j = 0$). Obviously, this condition can be checked in $\mathrm{O}(n|T||F|)$ time. □

Let us turn to problem CE($\mathcal{C}_{\mathrm{all}}$).

Observation 1. *If* $|AS(a)| > 0$ *holds for all* $a \in \widetilde{T} \cup \widetilde{F}$*, then* $(\widetilde{T}, \widetilde{F})$ *always has a consistent extension* f*. In other words, problem* CE($\mathcal{C}_{\mathrm{all}}$) *can be trivially solved.*

Proof. Let us consider an assignment $\alpha \in \mathbb{B}^{AS}$ such that $|ON(a^\alpha)|$ is odd for all $a \in \widetilde{T}$, and $|ON(b^\alpha)|$ is even for all $b \in \widetilde{F}$. (Since every vector contains at least one missing bit, we have such an assignment α.) Let f be the parity function, i.e., for which $f(v) = 1$ if and only if $|ON(v)|$ is odd. Then f is clearly a consistent extension of $(\widetilde{T}, \widetilde{F})$. □

Problem CE($\mathcal{C}_{\text{all}}$) becomes more complicated when not all input vectors have missing bits, although it remains polynomially solvable if each input vector contains at most one missing bit.

Theorem 4. *Problem* CE($\mathcal{C}_{\text{all}}$) *can be solved in polynomial time for a pBmb* $(\widetilde{T}, \widetilde{F})$ *for which every* $a \in \widetilde{T} \cup \widetilde{F}$ *satisfies* $|AS(a)| \leqslant 1$.

Proof. By definition, $(\widetilde{T}, \widetilde{F})$ has a consistent extension if and only if there exists an $\alpha \in \mathbb{B}^{AS}$ such that $\widetilde{T}^{\alpha} \cap \widetilde{F}^{\alpha} = \emptyset$. The latter condition is broken into the following two cases. For every pair of $a \in \widetilde{T}$ and $b \in \widetilde{F}$, (i) if $a, b \in \mathbb{B}^n$ (i.e., a, b contain no missing bits), then $a \neq b$ must hold, and (ii) if either a or b (or both) contains a missing bit and $a \approx b$, then $a^{\alpha} \neq b^{\alpha}$ must hold. Condition (i) is easy to check, and (ii) can be turned into a set of quadratic Boolean equations for $\alpha \in \mathbb{B}^{AS}$, which will be explained below.

Let j_a denote the index of the $*$ in each vector $a \in \widetilde{T} \cup \widetilde{F}$ (i.e., $AS(a) = \{(a, j_a)\}$), if any. Then (ii) can equivalently be formulated as the existence of an assignment $\alpha \in \mathbb{B}^{AS}$ satisfying the conditions

$$\begin{aligned}
&\alpha(a, j_a) \neq b_{j_a} && \text{if } |AS(a)| = 1 \text{ and } AS(b) = \emptyset, && (8)\\
&\alpha(b, j_b) \neq a_{j_b} && \text{if } AS(a) = \emptyset \text{ and } |AS(b)| = 1, && (9)\\
&\alpha(a, j_a) \neq b_{j_a} \text{ or } \alpha(b, j_b) \neq a_{j_b} && \text{if } |AS(a)| = 1,\ |AS(b)| = 1 \text{ and } j_a \neq j_b, && (10)\\
&\alpha(a, j_a) \neq \alpha(b, j_b) && \text{if } |AS(a)| = 1,\ |AS(b)| = 1 \text{ and } j_a = j_b, && (11)
\end{aligned}$$

for every pair of $a \in \widetilde{T}$ and $b \in \widetilde{F}$ with $a \approx b$. Obviously, condition (i) can be checked in $O(n|\widetilde{T}||\widetilde{F}|)$ time. To check (ii), let us observe that each of the conditions (8)–(11) can equivalently be represented as clauses in the variables $\alpha(v, j)$ for $(v, j) \in AS$. Namely, (8) and (9) can be represented by unit (or linear) clauses, (10) by a clause containing two variables, and (11) by the conjunction of two clauses, each of which contains two variables. For example, (11) is equivalent with the condition

$$1 = (\alpha(a, j_a) \vee \alpha(b, j_b))\big(\overline{\alpha(a, j_a)} \vee \overline{\alpha(b, j_b)}\big).$$

In total, we have a 2-SAT problem containing at most $2|\widetilde{T}||\widetilde{F}|$ clauses, which is solvable in time linear in its input size (see, e.g., [4]). This shows that problem CE($\mathcal{C}_{all}$) can be solved in $O(n|\widetilde{T}||\widetilde{F}|)$ time. □

Example 1. Consider $\widetilde{T}, \widetilde{F} \subseteq \{0, 1\}^3$ such that

$$\widetilde{T} = \left\{ \begin{array}{l} a^{(1)} = (1, 1, *) \\ a^{(2)} = (0, 0, 1) \\ a^{(3)} = (0, 1, *) \\ a^{(4)} = (*, 0, 0) \end{array} \right\}, \qquad \widetilde{F} = \left\{ \begin{array}{l} b^{(1)} = (1, 1, 1) \\ b^{(2)} = (0, *, 1) \\ b^{(3)} = (*, 0, 0) \end{array} \right\}.$$

Then we have the following 2-SAT:

$$\overline{\alpha(a^{(1)},3)}\alpha\big(b^{(2)},2\big)\big(\overline{\alpha(a^{(3)},3)}\vee\overline{\alpha(b^{(2)},2)}\big)\big(\alpha(a^{(4)},1)\vee\alpha(b^{(3)},1)\big)$$
$$\big(\overline{\alpha(a^{(4)},1)}\vee\overline{\alpha(b^{(3)},1)}\big)=1.$$

For this equation, the assignment $\alpha\in\mathbb{B}^{AS}$ given by $\alpha(a^{(1)},3)=\alpha(a^{(3)},3)=\alpha(a^{(4)},1)=0$ and $\alpha(b^{(2)},2)=\alpha(b^{(3)},1)=1$, is a satisfying solution.

In general, however, we have the following negative result.

Theorem 5. *Problem* CE($\mathcal{C}_{\text{all}}$) *is* NP-*complete, even if* $|AS(a)|\leqslant 2$ *holds for all* $a\in\widetilde{T}\cup\widetilde{F}$.

The highly technical proof is included in Appendix A.

3.3. Positive extensions

Let us consider subclasses of positive functions $\mathcal{C}\subseteq\mathcal{C}^+$. We shall see that all the cases of positive functions unresolved in Section 3.1 can be derived from the results about EXTENSION in [11] with the help of the following lemmas.

Recall that $a\in\mathbb{M}^n$ (i.e., the set $\{a^\alpha\mid\alpha\in AS(a)\}$) represents a subhypercube of $\mathbb{B}^n$. The following two lemmas show that $(\widetilde{T},\widetilde{F})$ has a robust extension f in $\mathcal{C}\subseteq\mathcal{C}^+$ if and only if f classifies the bottom element of each $a\in\widetilde{T}$ and the top element of each $b\in\widetilde{F}$ into 1 and 0, respectively, and that $(\widetilde{T},\widetilde{F})$ has a consistent extension f in $\mathcal{C}\subseteq\mathcal{C}^+$ if and only if f classifies the top element of each $a\in\widetilde{T}$ and the bottom element of each $b\in\widetilde{F}$ into 1 and 0, respectively.

Lemma 1. *Consider a class of functions* $\mathcal{C}\subseteq\mathcal{C}^+$. *For a pBmb* $(\widetilde{T},\widetilde{F})$, *let us associate a pdBf* (T^-,F^+) *by defining*

$$T^-=\big\{a^{\mathbf{0}}\mid a\in\widetilde{T}\big\},\qquad F^+=\big\{b^{\mathbf{1}}\mid b\in\widetilde{F}\big\},$$

where $\mathbf{0}\in\mathbb{B}^{AS(v)}$ *(respectively,* $\mathbf{1}\in\mathbb{B}^{AS(v)}$*) denotes the assignment of* 0*'s (respectively,* 1*'s) to all* $(v,i)\in AS(v)$. *Then, the pBmb* $(\widetilde{T},\widetilde{F})$ *has a robust extension in the class* $\mathcal{C}$ *if and only if the pdBf* (T^-,F^+) *has an extension in class* $\mathcal{C}$.

Proof. Let us assume first that the pBmb $(\widetilde{T},\widetilde{F})$ has a robust extension $f\in\mathcal{C}$. Then, by definition, f is an extension of the pdBf (T^-,F^+). For the converse direction, let us assume that the pdBf (T^-,F^+) has an extension g in class $\mathcal{C}$. For any assignment $\beta\in\mathbb{B}^{AS}$ and $a\in\widetilde{T}$, the vector $a^{\mathbf{0}}\in T^-$ satisfies $a^{\mathbf{0}}\leqslant a^\beta$, and hence $g(a^\beta)=1$ is implied by $g(a^\beta)\geqslant g(a^{\mathbf{0}})=1$. Similarly, for any assignment $\beta\in\mathbb{B}^{AS}$ and $b\in\widetilde{F}$, the vector $b^{\mathbf{1}}\in F^+$ satisfies $b^{\mathbf{1}}\geqslant b^\beta$, and hence $g(b^\beta)=0$ follows analogously. Therefore, g is a robust extension of the pBmb $(\widetilde{T},\widetilde{F})$ in the class $\mathcal{C}$. □

Lemma 2. *Consider a class of functions* $\mathcal{C}\subseteq\mathcal{C}^+$. *For a pBmb* $(\widetilde{T},\widetilde{F})$, *let us associate the pdBf* (T^+,F^-) *defined by*

$$T^+=\big\{a^{\mathbf{1}}\mid a\in\widetilde{T}\big\},\qquad F^-=\big\{b^{\mathbf{0}}\mid b\in\widetilde{F}\big\}.$$

Then, the pBmb $(\widetilde{T}, \widetilde{F})$ has a consistent extension in the class $\mathcal{C}$ if and only if the pdBf (T^+, F^-) has an extension in the same class.

Proof. Let us assume first that the pBmb $(\widetilde{T}, \widetilde{F})$ has a consistent extension $f \in \mathcal{C}$, i.e., that there exists an assignment $\beta \in \mathbb{B}^{AS}$ such that f is an extension of the pdBf $(\widetilde{T}^\beta, \widetilde{F}^\beta)$. Since $\mathcal{C} \subseteq \mathcal{C}^+$, for any $a \in \widetilde{T}$ (respectively, $b \in \widetilde{F}$), $f(a^\beta) = 1$ (respectively, $f(b^\beta) = 0$) implies $f(a^1) = 1$ (respectively, $f(b^0) = 0$) by $a^\beta \leqslant a^1$ (respectively, $b^\beta \geqslant b^0$). This implies that f is also an extension of the pdBf (T^+, F^-). The converse direction is immediate, since $(T^+, F^-) = (\widetilde{T}^\alpha, \widetilde{F}^\alpha)$ for the assignment $\alpha \in \mathbb{B}^{AS}$ defined by $\alpha(v, i) = 1$ if $v \in \widetilde{T}$, $(v, i) \in AS(v)$, and $\alpha(u, j) = 0$ if $u \in \widetilde{F}$, $(u, j) \in AS(u)$. □

The following theorem and its corollary directly follow from Lemmas 1 and 2, and from the results in [11].

Theorem 6. *If $\mathcal{C} \subseteq \mathcal{C}^+$ and problem* EXTENSION($\mathcal{C}$) *can be solved in polynomial time, then problems* RE($\mathcal{C}$) *and* CE($\mathcal{C}$) *can also be solved in polynomial time.*

Corollary 3. *Problems* RE($\mathcal{C}$) *and* CE($\mathcal{C}$) *can be solved in polynomial time for the following classes of functions $\mathcal{C}$:* (1) *positive,* (2) *positive $g(S_0, h_1(S_1))$-decomposable, and* (3) *positive k-DNF with fixed k.*

3.4. Threshold and Horn extensions

Let us consider the classes $\mathcal{C}_{\mathrm{TH}}$ of threshold, and $\mathcal{C}_{\mathrm{Horn}}$ of Horn functions. For these two classes, we shall see below that, although RE($\mathcal{C}$) is polynomially solvable, CE($\mathcal{C}$) is NP-complete even if $|AS(a)| \leqslant 1$ is assumed for all $a \in \widetilde{T} \cup \widetilde{F}$. For the related results on the existence of threshold and Horn extensions of a pdBf see, e.g., [11,24,26].

Theorem 7. *Problem* RE($\mathcal{C}_{\mathrm{TH}}$) *can be solved in polynomial time.*

Proof. For a pBmb $(\widetilde{T}, \widetilde{F})$, where $\widetilde{T}, \widetilde{F} \subseteq \mathbb{M}^n$, let us consider the following linear programming (LP) problem:

$$\begin{aligned}
&\max \xi = \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} z_i \\
&\text{subject to} \\
&\qquad \sum_{i \in ON(a)} w_i + \sum_{(a,i) \in AS(a)} y_i \geqslant t \qquad && \forall a \in \widetilde{T}, \\
&\qquad \sum_{i \in ON(b)} w_i + \sum_{(b,i) \in AS(b)} z_i \leqslant t - 1 && \forall b \in \widetilde{F}, \\
&\qquad y_i \leqslant w_i, \quad y_i \leqslant 0 && i = 1, 2, \ldots, n, \\
&\qquad z_i \geqslant w_i, \quad z_i \geqslant 0 && i = 1, 2, \ldots, n.
\end{aligned} \tag{12}$$

We claim that the LP problem (12) has a feasible solution with a finite optimum value ξ if and only if $(\widetilde{T}, \widetilde{F})$ has a robust extension in $\mathcal{C}_{\mathrm{TH}}$.

Informally, maximizing the objective function together with the last two constraints of (12) implies that $y_i = \min\{0, w_i\}$ and $z_i = \max\{0, w_i\}$ hold in the optimal solution. Therefore, this forces the left hand sides of the first two constraints in (12) to take their lowest possible values $\min_{\alpha \in AS(a)} \sum_{i=1}^n w_i (a^\alpha)_i$ for all $a \in \widetilde{T}$, and their highest possible values $\max_{\beta \in AS(b)} \sum_{i=1}^n w_i (b^\beta)_i$ for all $b \in \widetilde{F}$, respectively.

Let us assume first that $(\widetilde{T}, \widetilde{F})$ has a robust extension $f \in \mathcal{C}_{\mathrm{TH}}$, and let w_i, $i = 1, 2, \ldots, n$, and t be the coefficients in a realization (2). Then, since $\min\{w_i a_i^\alpha \mid \alpha \in \mathbb{B}\} = \min\{0, w_i\}$ and $\max\{w_i a_i^\alpha \mid \alpha \in \mathbb{B}\} = \max\{0, w_i\}$ hold for all i, $y_i = \min\{0, w_i\}$ and $z_i = \max\{0, w_i\}$ give a feasible solution to (12). The objective value satisfies $\xi = \sum_{i=1}^n y_i - \sum_{i=1}^n z_i \leqslant 0$, and (12) has a finite optimum.

Conversely, assume that w_i, y_i, z_i, $i = 1, 2, \ldots, n$, and t are an optimal solution of problem (12) (with a finite optimum). Then $y_i = \min\{0, w_i\}$ and $z_i = \max\{0, w_i\}$ hold since otherwise it could not be an optimum. This implies that w_i, $i = 1, 2, \ldots, n$, and t are finite, and define a threshold function that is a robust extension of $(\widetilde{T}, \widetilde{F})$. $\square$

Theorem 8. *Problem* $\mathrm{CE}(\mathcal{C}_{\mathrm{TH}})$ *is* NP-*complete, even if* $|AS(a)| \leqslant 1$ *holds for all* $a \in \widetilde{T} \cup \widetilde{F}$.

Proof. This problem is obviously in NP. To show its NP-hardness, consider a cubic CNF

$$\Phi = \bigwedge_{k=1}^{m} C_k, \qquad C_k = (u_k \vee v_k \vee w_k),$$

where u_k, v_k and w_k for $k = 1, 2, \ldots, m$ are literals from set $L = \{x_1, \bar{x}_1, \ldots, x_n, \bar{x}_n\}$. Let $L' = \{x_1', \bar{x}_1', \ldots, x_n', \bar{x}_n'\}$, and define $\widetilde{T}, \widetilde{F} \subseteq \mathbb{M}^{L \cup L'}$ as follows.

$$\begin{aligned}
\widetilde{T} &= \{a^{x_i} = (\{x_i'\}; \{x_i\}), a^{\bar{x}_i} = (\{\bar{x}_i'\}; \{\bar{x}_i\}) \mid i = 1, 2, \ldots, n\}, \\
\widetilde{F} &= \{b^{(0)} = (\emptyset; \emptyset)\} \cup \{(b^{x_i} = \{x_i, \bar{x}_i\}; \emptyset), b^{x_i'} = (\{x_i', \bar{x}_i'\}; \emptyset), \mid i = 1, 2, \ldots, n\} \\
&\qquad \cup \{b^{C_k} = (\{u_k, v_k, w_k, u_k', v_k', w_k'\}; \emptyset) \mid k = 1, 2, \ldots, m\},
\end{aligned} \tag{13}$$

where $(R; S)$ denotes the vector $v \in \mathbb{M}^{L \cup L'}$ such that $ON(v) = R$ and $AS(v) = \{(v, j) \mid j \in S\}$. It is easy to see that $|AS(a)| \leqslant 1$ holds for all $a \in \widetilde{T} \cup \widetilde{F}$. We claim that this pBmb $(\widetilde{T}, \widetilde{F})$ has a consistent threshold extension if and only if the 3-SAT problem for the CNF Φ has a solution, which will complete the proof.

Let us assume that $\alpha \in \mathbb{B}^{AS}$ is an assignment such that $(\widetilde{T}^\alpha, \widetilde{F}^\alpha)$ has a threshold extension:

$$f(d) = \begin{cases} 1 & \text{if } \sum_{z \in L \cup L'} w_z d_z \geqslant t, \\ 0 & \text{if } \sum_{z \in L \cup L'} w_z d_z \leqslant t - 1, \end{cases}$$

where $d \in \mathbb{B}^{L \cup L'}$. Note first that $t \geqslant 1$ follows from $f(b^{(0)}) = 0$ and $\sum_{z \in L \cup L'} w_z b_z^{(0)} = 0$. We shall show that $\alpha(a^{x_i}, x_i) \neq \alpha(a^{\bar{x}_i}, \bar{x}_i)$ must hold for (a^{x_i}, x_i), $(a^{\bar{x}_i}, \bar{x}_i) \in AS$. If $\alpha(a^{x_i}, x_i) = \alpha(a^{\bar{x}_i}, \bar{x}_i) = 1$ holds, then $(a^{x_i})^\alpha \in \widetilde{T}^\alpha$ and $(a^{\bar{x}_i})^\alpha \in \widetilde{T}^\alpha$, respectively, imply $w_{x_i} + w_{x_i'} \geqslant t$ and $w_{\bar{x}_i} + w_{\bar{x}_i'} \geqslant t$, and hence

$$w_{x_i} + w_{x_i'} + w_{\bar{x}_i} + w_{\bar{x}_i'} \geqslant 2t. \tag{14}$$

However, $b^{x_i} \in \widetilde{F}^\alpha$ and $b^{x_i'} \in \widetilde{F}^\alpha$, respectively, imply $w_{x_i} + w_{\bar{x}_i} < t$ and $w_{x_i'} + w_{\bar{x}_i'} < t$, and hence $w_{x_i} + w_{x_i'} + w_{\bar{x}_i} + w_{\bar{x}_i'} < 2t$, which is a contradiction to (14). Furthermore, if $\alpha(a^{x_i}, x_i) = \alpha(a^{\bar{x}_i}, \bar{x}_i) = 0$ holds, then $(a^{x_i})^\alpha \in \widetilde{T}^\alpha$ and $(a^{\bar{x}_i})^\alpha \in \widetilde{T}^\alpha$, respectively, imply $w_{x_i'} \geqslant t$ and $w_{\bar{x}_i'} \geqslant t$, and hence $w_{x_i'} + w_{\bar{x}_i'} \geqslant 2t > t$ (by $t > 0$), which is a contradiction to $f(b^{x_i'}) = 0$. Hence $\alpha(a^{x_i}, x_i) \neq \alpha(a^{\bar{x}_i}, \bar{x}_i)$ holds. Let us now define a binary vector $x^* \in \mathbb{B}^n$ by

$$x_i^* = \begin{cases} 1 & \text{if } \alpha(a^{x_i}, x_i) = 0, \\ 0 & \text{otherwise,} \end{cases}$$

and show that this y satisfies $\Phi(x^*) = 1$. For this, assume otherwise that there is a clause C_k that satisfies $C_k(x^*) = 0$ (i.e., $u_k = v_k = w_k = 0$ holds by x^*), that is, $\alpha(a^{u_k}, u_k) = \alpha(a^{v_k}, v_k) = \alpha(a^{w_k}, w_k) = 1$. Then taking three vectors $(a^{u_k})^\alpha, (a^{v_k})^\alpha, (a^{w_k})^\alpha \in \widetilde{T}^\alpha$, we have $w_{u_k} + w_{u_k'} \geqslant t$, $w_{v_k} + w_{v_k'} \geqslant t$ and $w_{w_k} + w_{w_k'} \geqslant t$, and hence $w_{u_k} + w_{v_k} + w_{w_k} + w_{u_k'} + w_{v_k'} + w_{w_k'} \geqslant 3t > t$, which is a contradiction to $f(b^{C_k}) = 0$.

For the converse direction, let us assume that $\Phi(x^*) = 1$ holds for some $x^* \in \mathbb{B}^n$. Let us define an assignment $\alpha \in \mathbb{B}^{AS}$ by $\alpha(a^{x_i}, x_i) = \bar{x}_i^*$ and $\alpha(a^{\bar{x}_i}, \bar{x}_i) = x_i^*$ for $i = 1, 2, \ldots, n$, and let

$$w_z = \begin{cases} -3 & \text{if either } z = x_i \text{ and } x_i^* = 1, \text{ or } z = \bar{x}_i \text{ and } x_i^* = 0, \\ +2 & \text{if either } z = x_i \text{ and } x_i^* = 0, \text{ or } z = \bar{x}_i \text{ and } x_i^* = 1, \\ +1 & \text{if either } z = x_i' \text{ and } x_i^* = 1, \text{ or } z = \bar{x}_i' \text{ and } x_i^* = 0, \\ -1 & \text{if either } z = x_i' \text{ and } x_i^* = 0, \text{ or } z = \bar{x}_i' \text{ and } x_i^* = 1, \end{cases}$$

and $t = 1$. Then $\sum_{z \in L \cup L'} w_z a_z \geqslant 1$ holds for all $a \in \widetilde{T}^\alpha$, and $\sum_{z \in L \cup L'} w_z b_z \leqslant 0$ holds for all $b \in \widetilde{F}^\alpha$. Hence $(\widetilde{T}, \widetilde{F})$ has a consistent threshold extension. □

Theorem 9. *Problem* RE($\mathcal{C}_{\text{Horn}}$) *can be solved in polynomial time.*

Proof. Let $(\widetilde{T}, \widetilde{F})$ be a pBmb. For each $a \in \widetilde{T}$, let us define $B(a) = \{b \in \widetilde{F} \mid b \gtrapprox a\}$. We claim that $(\widetilde{T}, \widetilde{F})$ has a robust Horn extension if and only if for every $a \in \widetilde{T}$, there exists an index j such that $a_j = 0$ and $b_j = 1$ for all $b \in B(a)$. The latter condition can be easily checked in $\mathrm{O}(n|\widetilde{T}||\widetilde{F}|)$ time.

To prove the claim, let us assume first that, for every $a \in \widetilde{T}$, there exists an index j such that $a_j = 0$ and $b_j = 1$ for all $b \in B(a)$. Then for any $\alpha \in \mathbb{B}^{AS}$, all $b^\alpha \in B(a)^\alpha$ satisfies $b_j^\alpha = 1$. Thus, for the Horn term

$$t_a = \Bigl(\prod_{i \in ON(a)} x_i\Bigr)\bar{x}_j,$$

we have $t_a(a^\alpha) = 1$ and $t_a(b^\alpha) = 0$ for all $\alpha \in \mathbb{B}^{AS}$ and $b \in \widetilde{F}$. Hence, the Horn DNF

$$\varphi = \bigvee_{a \in \widetilde{T}} t_a$$

provides a Horn extension of $(\widetilde{T}, \widetilde{F})$.

For the converse direction, let us assume that for some $a \in \widetilde{T}$, every index j with $a_j = 0$ has a vector $b \in B(a)$ with $b_j \in \{0, *\}$. For such a vector a, consider the assignments $\alpha \in \mathbb{B}^{AS(a)} \cup \mathbb{B}^{AS(B(a))}$ defined by

$$\alpha(a, i) = \begin{cases} \prod_{b \in B(a) \text{ s.t. } b_i \neq *} b_i & \text{if there is a vector } b \in B(a) \text{ with } b_i \in \{0, 1\}, \\ 1 & \text{otherwise,} \end{cases}$$

for $(a, i) \in AS(a)$, and $\alpha(b, i) = a_i^\alpha$ for $(b, i) \in AS(B(a))$. Then $\{b^\alpha \in \widetilde{F}^\alpha \mid b^\alpha \geqslant a^\alpha\} = B(a)^\alpha$ satisfies

$$a^\alpha = \bigwedge_{\{b^\alpha \in \widetilde{F}^\alpha \mid b^\alpha \geqslant a^\alpha\}} b^\alpha,$$

by the above assumption on a and $B(a)$, where $\wedge$ denotes the component-wise AND operation, e.g., $(010111) \wedge (100101) = (000101)$. However, it is known [11,26] that a pdBf (T, F) has an extension in $\mathcal{C}_{\text{Horn}}$ if and only if

$$a \neq \bigwedge_{b \in F \text{ s.t. } b \geqslant a} b \tag{15}$$

holds for all $a \in T$. Hence, $(\widetilde{T}^\alpha, \widetilde{F}^\alpha)$ has no extension in $\mathcal{C}_{\text{Horn}}$. $\square$

Theorem 10. *Problem* CE($\mathcal{C}_{\text{Horn}}$) *is* NP-*complete, even if* $|AS(a)| \leqslant 1$ *holds for all* $a \in \widetilde{T} \cup \widetilde{F}$.

Proof. It can be proved by a reduction from the 2-coloring problem of 3-uniform hypergraphs (for the details see [12]). $\square$

3.5. Decomposable extensions

Let us consider two basic classes of decomposable functions, $\mathcal{C}_{G(S_0, G(S_1))}$, and $\mathcal{C}^+_{G(S_0, G(S_1))}$, where $S_0, S_1 \subseteq V$. It is known that EXTENSION($\mathcal{C}_{G(S_0, G(S_1))}$) and EXTENSION($\mathcal{C}^+_{G(S_0, G(S_1))}$) can be solved in polynomial time [8].

Let us first consider RE($\mathcal{C}_{G(S_0, G(S_1))}$), where it is emphasized that $S_0 \cap S_1 \neq \emptyset$ generally holds. For a subset $\widetilde{S} \subseteq \mathbb{M}^V$, let $AS_k(\widetilde{S}) = \{(v, j) \in AS(\widetilde{S}) \mid j \in S_k\}$ for $k = 0, 1$, and $AS_k = AS_k(\widetilde{T} \cup \widetilde{F})$. For a given pBmb $(\widetilde{T}, \widetilde{F})$, we define a graph $G_{(\widetilde{T}, \widetilde{F})} = (W, E_1 \cup E_2)$ by

$$\begin{aligned} W = \big\{w, w' \mid & \text{there exist } a \in \widetilde{T}, b \in \widetilde{F}, \alpha \in \mathbb{B}^{AS_0(a)} \text{ and } \beta \in \mathbb{B}^{AS_0(b)} \\ & \text{such that } a^\alpha[S_0] = b^\beta[S_0] \text{ and } w = a^\alpha[S_1], w' = b^\beta[S_1]\big\} \\ & (\subseteq \mathbb{B}^{S_1 \cap S_0} \times \mathbb{M}^{S_1 \setminus S_0}) \end{aligned}$$

$$\begin{aligned} E_1 = \big\{(w, w') \mid & \text{there exist } a \in \widetilde{T}, b \in \widetilde{F}, \alpha \in \mathbb{B}^{AS_0(a)} \text{ and } \beta \in \mathbb{B}^{AS_0(b)} \\ & \text{such that } a^\alpha[S_0] = b^\beta[S_0] \text{ and } w = a^\alpha[S_1], w' = b^\beta[S_1]\big\} \end{aligned}$$

$$\begin{aligned} E_2 = \big\{(u, v) \mid & u, v \in W \text{ with } u = a^\alpha[S_1] \text{ and } v = b^\beta[S_1] \text{ for } a, b \in \widetilde{T} \cup \widetilde{F} \text{ and} \\ & \alpha \in \mathbb{B}^{AS_0(a)}, \beta \in \mathbb{B}^{AS_0(b)}, \text{ for which } a^\alpha[S_1] \approx b^\beta[S_1] \text{ holds}\big\}. \end{aligned}$$

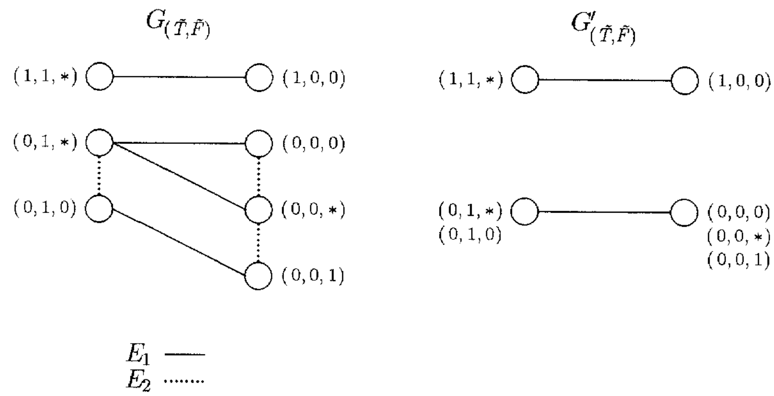

Fig. 1. Graphs $G_{(\widetilde{T},\widetilde{F})}$ and $G'_{(\widetilde{T},\widetilde{F})}$ of $(\widetilde{T},\widetilde{F})$ in Example 2.

Furthermore, denote by $G'_{(\widetilde{T},\widetilde{F})}$ the graph obtained from $G_{(\widetilde{T},\widetilde{F})}$ by contracting all edges in E_2. Note that $G'_{(\widetilde{T},\widetilde{F})}$ has self-loops if there are edges (w, w') in $E_1 \cap E_2$, and any graph containing a self-loop is not bipartite.

Example 2. Let $S_0 = \{1,2,3,4\}$, $S_1 = \{4,5,6\}$ and $V = S_0 \cup S_1$ (i.e., $V = \{1,2,\ldots,6\}$), and define $\widetilde{T}, \widetilde{F} \subseteq \{0,1\}^V$ by

$$\widetilde{T} = \left\{ \begin{array}{l} a^{(1)} = (1,1,1,1,*,0) \\ a^{(2)} = (0,*,1,*,1,*) \\ a^{(3)} = (0,0,0,0,1,0) \end{array} \right\}, \qquad \widetilde{F} = \left\{ \begin{array}{l} b^{(1)} = (0,1,*,*,0,0) \\ b^{(2)} = (0,0,1,0,0,*) \\ b^{(3)} = (0,0,0,0,0,1) \end{array} \right\}.$$

Graphs $G_{(\widetilde{T},\widetilde{F})}$ and $G'_{(\widetilde{T},\widetilde{F})}$ are given in Fig. 1, where solid edges stand for E_1 and dotted edges for E_2. Note that, for example, $G_{(\widetilde{T},\widetilde{F})}$ does not have vertex $(1,*,0) \in \mathbb{B}^{S_1\cap S_0} \times \mathbb{M}^{S_1\setminus S_0}$, which is obtained from $a^{(1)} \in \widetilde{T}$, since there is no $b \in \widetilde{F}$ such that $(a^{(1)})^{\alpha}[S_0] = b^{\beta}[S_0]$ holds for some $\alpha \in \mathbb{B}^{AS_0(a^{(1)})}$ and $\beta \in \mathbb{B}^{AS_0(b)}$.

Lemma 3. *Let $(\widetilde{T},\widetilde{F})$ be a pBmb. Then $(\widetilde{T},\widetilde{F})$ has a robust $G(S_0, G(S_1))$-decomposable extension if and only if $G'_{(\widetilde{T},\widetilde{F})}$ is bipartite.*

Proof. Let us first show the only-if part. Assume that $(\widetilde{T},\widetilde{F})$ has a robust $G(S_0, G(S_1))$-decomposable extension, but $G'_{(\widetilde{T},\widetilde{F})}$ is not bipartite. In other words, there is a cycle $w^{(0)} \to w^{(1)} \to \cdots \to w^{(l)} (= w^{(0)})$ in $G_{(\widetilde{T},\widetilde{F})} = (W, E_1 \cup E_2)$ such that

$$\left|E_1 \cap \left\{(w^{(i)}, w^{(i+1)}) \mid i = 0,1,\ldots,l-1\right\}\right| \text{ is odd.} \tag{16}$$

Let us consider the values of h_1 on $\{(w^{(i)})^{\alpha} \mid \alpha \in \mathbb{B}^{AS_1(w^{(i)})}\}$, $i = 0,1,\ldots,l-1$. For each $(w^{(i)}, w^{(i+1)}) \in E_1$, by the definitions of E_1 and a robust extension, we must have

$h_1((w^{(i)})^\beta) \neq h_1((w^{(i+1)})^\gamma)$ for all assignments $\beta \in \mathbb{B}^{AS_1(w^{(i)})}$ and $\gamma \in \mathbb{B}^{AS_1(w^{(i+1)})}$. This means that

$$\begin{aligned} h_1\big((w^{(i)})^\beta\big) &= p \qquad \text{for all } \beta \in \mathbb{B}^{AS_1(w^{(i)})}, \\ h_1\big((w^{(i+1)})^\gamma\big) &= \bar{p} \qquad \text{for all } \gamma \in \mathbb{B}^{AS_1(w^{(i+1)})}, \end{aligned} \tag{17}$$

for some $p \in \{0, 1\}$. On the other hand, if $(w^{(i)}, w^{(i+1)}) \in E_2$,

$$h_1\big((w^{(i)})^\beta\big) = h_1\big((w^{(i+1)})^\gamma\big) \tag{18}$$

holds for all $\beta \in \mathbb{B}^{AS_1(w^{(i)})}$ and $\gamma \in \mathbb{B}^{AS_1(w^{(i+1)})}$, because the definition of W and (17) imply that $h_1\big((w^{(i)})^\beta\big) = p$ for all $\beta \in \mathbb{B}^{AS_1(w^{(i)})}$, $h_1\big((w^{(i+1)})^\gamma\big) = q$ for all $\gamma \in \mathbb{B}^{AS_1(w^{(i+1)})}$, and $p = q$ by $(w^{(i)}, w^{(i+1)}) \in E_2$. Thus (17) and (18) contradict (16).

Conversely, if $G'_{(\widetilde{T},\widetilde{F})}$ is bipartite, then there is a partition $(Y,\ W \setminus Y)$ of W such that

$$\begin{aligned} &E_1 \subseteq Y \times (W \setminus Y), \\ &E_2 \subseteq (Y \times Y) \cup \big((W \setminus Y) \times (W \setminus Y)\big). \end{aligned} \tag{19}$$

By (19), we can define the value of h_1 for W by

$$h_1\big((w)^\beta\big) = \begin{cases} 1 & \text{if } w \in Y \text{ and } \beta \in \mathbb{B}^{AS_1(w)}, \\ 0 & \text{if } w \in W \setminus Y \text{ and } \beta \in \mathbb{B}^{AS_1(w)}. \end{cases} \tag{20}$$

The h_1 values for other vectors $(v)^\gamma$ are determined arbitrarily, where $v \in (\widetilde{T} \cup \widetilde{F})[S_1] \setminus W$ and $\gamma \in AS_1(v)$. Furthermore, define g by

$$\begin{aligned} g\big(a^\alpha[S_0], h_1(a^\alpha[S_1])\big) &= 1 \qquad \text{for all } a \in \widetilde{T} \text{ and } \alpha \in \mathbb{B}^{AS(a)}, \\ g\big(b^\alpha[S_0], h_1(b^\alpha[S_1])\big) &= 0 \qquad \text{for all } b \in \widetilde{F} \text{ and } \alpha \in \mathbb{B}^{AS(b)}. \end{aligned}$$

If $a^\alpha[S_0] = b^\alpha[S_0]$ holds for some $a \in \widetilde{T}$, $b \in \widetilde{F}$ and $\alpha \in \mathbb{B}^{AS(\{a,b\})}$, then $w_a, w_b \in W$ with $w_a \approx a^\alpha[S_1]$ and $w_b \approx b^\alpha[S_1]$ satisfy $(w_a, w_b) \in E_1$, and we have $h_1(a^\alpha[S_1]) \neq h_1(b^\alpha[S_1])$ by (19) and (20). Hence g is also well-defined. Therefore, we see that $(\widetilde{T}, \widetilde{F})$ has a robust $G(S_0, G(S_1))$-decomposable extension. $\square$

In general, however, the size of graph $G'_{(\widetilde{T},\widetilde{F})}$ can be exponential in $|S_0|$, and the above lemma does not lead to an efficient algorithm of $\mathrm{RE}(\mathcal{C}_{G(S_0,G(S_1))})$.

Theorem 11. *Problem* $\mathrm{RE}(\mathcal{C}_{G(S_0,G(S_1))})$ *is* co-NP-*complete.*

Proof. First we show that the problem is in co-NP. For a pBmb $(\widetilde{T}, \widetilde{F})$, we show that every simple cycle C in $G'_{(\widetilde{T},\widetilde{F})}$ satisfies $|C| \leqslant |\widetilde{T}| + |\widetilde{F}|$. By the definition of E_1 and E_2, $w[S_0 \cap S_1] = w'[S_0 \cap S_1]$ holds for all edges $(w, w') \in E_1 \cup E_2$. Thus

$$w[S_0 \cap S_1] = w'[S_0 \cap S_1] \tag{21}$$

holds if there is a path from w to w' in $G_{(\widetilde{T},\widetilde{F})}$. In particular, all vertices in a cycle C in $G'_{(\widetilde{T},\widetilde{F})}$ have this property, and, by the definition of $G'_{(\widetilde{T},\widetilde{F})}$, all vertices w in C have different $w[S_1 \setminus S_0]$, implying that they are generated from different vectors in $\widetilde{T} \cup \widetilde{F}$. This proves $|C| \leqslant |\widetilde{T}| + |\widetilde{F}|$. Since $(\widetilde{T}, \widetilde{F})$ has a robust $G(S_0, G(S_1))$-decomposable extension if and only if there is no cycle C of odd length in $G'_{(\widetilde{T},\widetilde{F})}$, a negative answer for $\mathrm{RE}(\mathcal{C}_{G(S_0,G(S_1))})$ can be certified by an odd cycle of length no more than $|\widetilde{T}| + |\widetilde{F}|$, hence $\mathrm{RE}(\mathcal{C}_{G(S_0,G(S_1))})$ belongs to co-NP.

We next show its co-NP-hardness. Let $\mathcal{H} = (U, E)$ be a 3-uniform hypergraph, where $U = \{1, \ldots, n\}$, $E = \{H_i \mid i = 1, \ldots, m\}$, $|H_i| = 3$ for all i, and m is odd. Let $S_0 = \{1, 2, \ldots, n+m\}$, $S_1 = \{1, 2, \ldots, n\} \cup \{n+m+1, n+m+2, \ldots, n+2m\}$. Obviously, $V = S_0 \cup S_1 = \{1, 2, \ldots, n+2m\}$ and $U = S_0 \cap S_1 = \{1, 2, \ldots, n\}$ hold in this case. Then define $\widetilde{T}, \widetilde{F} \subseteq \mathbb{M}^V$ as follows.

$$\widetilde{T} = \{(A_i \cup \{n+i\} \cup \{n+m+i\}; U \setminus H_i) \mid A_i \subset H_i, A_i \neq \emptyset, i \in \{1, 2, \ldots, m\}\}$$
$$\widetilde{F} = \{(A_i \cup \{n+i\} \cup \{n+m+(i \pmod m)+1\}; U \setminus H_i) \mid A_i \subset H_i, A_i \neq \emptyset, \\ i \in \{1, 2, \ldots, m\}\},$$

where $\subset$ denotes the proper inclusion, and $(R; S)$ denotes the vector $v \in \mathbb{M}^V$ such that $ON(v) = R$ and $AS(v) = \{(v, i) \mid i \in S\}$. We claim that this pBmb $(\widetilde{T}, \widetilde{F})$ has a robust extension in $\mathcal{C}_{G(S_0,G(S_1))}$ if and only if $\mathcal{H}$ is not 2-colorable, where $\mathcal{H}$ is called 2-colorable if there exists a partition $(C, U \setminus C)$ of U such that $H_i \not\subseteq C$ and $H_i \not\subseteq U \setminus C$ holds for all $H_i \in E$. This completes the proof because deciding if $\mathcal{H}$ is 2-colorable is NP-complete (see, e.g., [21]), even if m is restricted to be odd. For this $G_{(\widetilde{T},\widetilde{F})} = (W, E_1 \cup E_2)$, we have $E_2 = \emptyset$, because any $(v, j) \in AS$ satisfies $j \in S_0 \cap S_1$, implying that $v \in \mathbb{B}^{S_1}$ holds for any vertex v in $G_{(\widetilde{T},\widetilde{F})}$. This means that $G'_{(\widetilde{T},\widetilde{F})}$ is bipartite if and only if so is $G_{(\widetilde{T},\widetilde{F})}$. Thus Lemma 3 tells that $(\widetilde{T}, \widetilde{F})$ has a robust $G(S_0, G(S_1))$-decomposable extension if and only if $G_{(\widetilde{T},\widetilde{F})}$ is bipartite.

Let us first assume that $(C, U \setminus C)$ is a 2-coloring of $\mathcal{H}$, i.e., $C \cap H_i \neq \emptyset$ and $(U \setminus C) \cap H_i \neq \emptyset$ for all $H_i \in E$. Then C can be represented by

$$C = \bigcup_{i=1}^{m} A_i^* \tag{22}$$

for some $\emptyset \neq A_i^* \subset H_i$, $i = 1, 2, \ldots, m$, and we have

$$((C \cup \{n+m+i\}; \emptyset), (C \cup \{n+m+(i \pmod m)+1\}; \emptyset)) \in E_1, \\ \text{for } i = 1, 2, \ldots, m.$$

Hence, we have a cycle $w^{(1)} \to w^{(2)} \to \cdots \to w^{(m)} \to w^{(1)}$ in $G_{(\widetilde{T},\widetilde{F})}$, where

$$w^{(i)} = (C \cup \{n+m+i\}; \emptyset), \quad i = 1, 2, \ldots, m.$$

Since m is chosen to be odd, this implies that $(\widetilde{T}, \widetilde{F})$ has no robust extension in $\mathcal{C}_{G(S_0,G(S_1))}$.

For the converse direction, let us assume that $G_{(\widetilde{T},\widetilde{F})}$ has a cycle. By property (21), we can assume that the cycle belongs to $G_{(\widetilde{T},\widetilde{F})}[W_C]$ for some $C \subset U$, where $W_C = \{w \in W \mid ON(w[U]) = C\}$ and $G_{(\widetilde{T},\widetilde{F})}[W_C]$ is the subgraph of $G_{(\widetilde{T},\widetilde{F})}$ induced by W_C. By the definition of the above $(\widetilde{T}, \widetilde{F})$, such a cycle must be of the form

$$\big(C \cup \{n+m+1\}; \emptyset\big) \to \big(C \cup \{n+m+2\}; \emptyset\big) \to \cdots \to \big(C \cup \{n+2m\}; \emptyset\big) \\ \to \big(C \cup \{n+m+1\}; \emptyset\big).$$

Thus the length of this cycle is odd. This C obviously satisfies (22) and is a 2-coloring of $\mathcal{H}$. □

However, we can point out an important special case in which $\mathrm{RE}(\mathcal{C}_{G(S_0,G(S_1))})$ is polynomially solvable.

Theorem 12. *If $S_0 \cap S_1 = \emptyset$, problem $\mathrm{RE}(\mathcal{C}_{G(S_0,G(S_1))})$ can be solved in polynomial time.*

Proof. In this case, graph $G_{(\widetilde{T},\widetilde{F})} = (W, E_1 \cup E_2)$ can be represented by

$$\begin{aligned} W &= \big\{a[S_1], b[S_1] \mid a \in \widetilde{T}, b \in \widetilde{F} \text{ and } a[S_0] \approx b[S_0]\big\}, \\ E_1 &= \big\{\big(a[S_1], b[S_1]\big) \mid a \in \widetilde{T}, b \in \widetilde{F} \text{ and } a[S_0] \approx b[S_0]\big\}, \\ E_2 &= \big\{\big(a[S_1], b[S_1]\big) \mid a[S_1], b[S_1] \in W \text{ and } a[S_1] \approx b[S_1]\big\}. \end{aligned}$$

It is easy to see that this graph $G'_{(\widetilde{T},\widetilde{F})}$ has polynomially many vertices and can be constructed in polynomial time. Then, by applying Lemma 3, $\mathrm{RE}(\mathcal{C}_{G(S_0,G(S_1))})$ can be solved in polynomial time. □

Similarly to other classes, $\mathrm{CE}(\mathcal{C}_{G(S_0,G(S_1))})$ appears more difficult than $\mathrm{RE}(\mathcal{C}_{G(S_0,G(S_1))})$.

Theorem 13. *Problem $\mathrm{CE}(\mathcal{C}_{G(S_0,G(S_1))})$ is NP-complete, even if $S_0 \cap S_1 = \emptyset$ and $|AS(a)| \leqslant 1$ for all $a \in \widetilde{T} \cup \widetilde{F}$.*

Proof. The proof can be done again by a reduction from the 2-colorability of 3-uniform hypergraphs [21]. For the details see [12]. □

Let us remark finally that problems $\mathrm{RE}(\mathcal{C}^+_{G(S_0,G(S_1))})$ and $\mathrm{CE}(\mathcal{C}^+_{G(S_0,G(S_1))})$ can be solved in polynomial time by Corollary 3.

3.6. *k-DNF extensions*

Let us consider the classes $\mathcal{C}_{k\text{-DNF}}$ and $\mathcal{C}^+_{k\text{-DNF}}$. Recall that a DNF

$$\varphi = \bigvee_{i=1}^{m} \Big(\prod_{j \in P_i} x_j \prod_{j \in N_i} \bar{x}_j \Big)$$

is a k-DNF if $|N_i \cup P_i| \leqslant k$ for $i = 1, \ldots, m$, and it is a *positive* k-DNF if, in addition, $N_i = \emptyset$ for $i = 1, \ldots, m$.

For general k (i.e., k is a parameter included in the problem specification), Corollary 1 tells that problems $\mathrm{CE}(\mathcal{C}_{k\text{-DNF}})$ and $\mathrm{CE}(\mathcal{C}^+_{k\text{-DNF}})$ are NP-complete, and problems $\mathrm{RE}(\mathcal{C}_{k\text{-DNF}})$ and $\mathrm{RE}(\mathcal{C}^+_{k\text{-DNF}})$ are NP-hard. However, for a fixed k, problems $\mathrm{RE}(\mathcal{C}^+_{k\text{-DNF}})$ and $\mathrm{CE}(\mathcal{C}^+_{k\text{-DNF}})$ can be solved in polynomial time, by Corollary 3.

Among the remaining problems, we start with problem RE($\mathcal{C}_{k\text{-DNF}}$) for a fixed k. For a vector $v \in \mathbb{M}^n$, let $A(v)$ denote the assignment to the variables x_i defined by

$$A(v) = (x_i \leftarrow v_i \mid v_i \neq *). \tag{23}$$

For example, if $v = (1, *, 0, 0, *)$, then $A(v) = (x_1 \leftarrow 1, x_3 \leftarrow 0, x_4 \leftarrow 0)$. Let $f_{A(v)}$ (respectively, $\varphi_{A(v)}$) denote the function (respectively, DNF) obtained by fixing the variables x_i as specified by $A(v)$.

Lemma 4. *Consider a vector $v \in \mathbb{M}^n$ and a term $t = \prod_{j\in P} x_j \prod_{j\in N} \bar{x}_j$. Then $t(v^\alpha) = 0$ holds for all assignments $\alpha \in \mathbb{B}^{AS(v)}$ if and only if $ON(v) \cap N \neq \emptyset$ or $OFF(v) \cap P \neq \emptyset$.*

Proof. It is easy to see that the if part holds. For the only-if part, assume $ON(v) \cap N = OFF(v) \cap P = \emptyset$, and define an assignment $\alpha \in \mathbb{B}^{AS(v)}$ by

$$\alpha(v,i) = \begin{cases} 1, & \text{if } i \in P, (v,i) \in AS(v), \\ 0, & \text{if } i \in N, (v,i) \in AS(v). \end{cases}$$

This assignment $\alpha \in \mathbb{B}^{AS(v)}$ obviously satisfies $t(v^\alpha) = 1$. □

Lemma 5. *Let φ be a* DNF *of n variables, and let $v \in \mathbb{M}^n$. For a subset $Q \subseteq AS(v)$ and an assignment $\alpha \in \mathbb{B}^Q$,*

(i) *$\varphi(v^{(\alpha,\beta)}) = 1$ holds for all assignments $\beta \in \mathbb{B}^{AS(v)\setminus Q}$ if and only if $\varphi_{A(v^\alpha)} = \top$, and*

(ii) *$\varphi(v^{(\alpha,\beta)}) = 0$ holds for all assignments $\beta \in \mathbb{B}^{AS(v)\setminus Q}$ if and only if $\varphi_{A(v^\alpha)} = \bot$,*

where (α, β) is the concatenation of assignments α and β.

Proof. Similar to the proof of Lemma 4. □

For a k-DNF φ, the problem of checking if $\varphi \neq \top$ is called k-NONTAUTOLOGY [21]. It is known that its complexity is the same as of k-SAT. For $k \leqslant 2$, k-SAT can be solved in polynomial time, but for $k \geqslant 3$, k-SAT is NP-complete [21]. The problem of checking if $\varphi = \top$ is called k-TAUTOLOGY. It follows from the result about k-SAT that k-TAUTOLOGY is solvable in polynomial time for $k \leqslant 2$ but is co-NP-complete for $k \geqslant 3$.

Theorem 14. *If $k \leqslant 2$, then problem* RE($\mathcal{C}_{k\text{-DNF}}$) *can be solved in polynomial time.*

Proof. The following algorithm solves problem RE($\mathcal{C}_{k\text{-DNF}}$).

Algorithm CHECK-RE($\mathcal{C}_{k\text{-DNF}}$)

Input: a pBmb $(\widetilde{T}, \widetilde{F})$, where $\widetilde{T}, \widetilde{F} \subseteq \mathbb{M}^n$.

Output: If a pBmb $(\widetilde{T}, \widetilde{F})$ has a robust extension in $\mathcal{C}_{k\text{-DNF}}$, then output such a DNF φ; otherwise, NO.

Step 1. Generate all possible terms with at most k literals. Let φ be the disjunction of all those terms t for which $t(b^\alpha) = 0$ holds for all $b \in \widetilde{F}$ and $\alpha \in \mathbb{B}^{AS(b)}$.

Step 2. If $\varphi_{A(a)} = \top$ for all $a \in \widetilde{T}$, then output φ; otherwise, output NO.

It is easy to see that the φ obtained in step 1 is a k-DNF, and furthermore it is the maximum k-DNF (with respect to $T(\varphi)$) such that $\varphi(b^\alpha) = 0$ for all $b \in \widetilde{F}$ and $\alpha \in \mathbb{B}^{AS(b)}$. By Lemma 5 (the case of $Q = \emptyset$), if φ passes the test of step 2, then $\varphi(a^\alpha) = 1$ must hold for all $a \in \widetilde{T}$ and $\alpha \in \mathbb{B}^{AS(a)}$. Hence this φ represents a robust extension of $(\widetilde{T}, \widetilde{F})$; otherwise there is no robust extension.

Let us next consider its time complexity. In Step 1, by Lemma 4, checking of each term t can be done in $\mathrm{O}(n|\widetilde{F}|)$ time. Since there are at most $M = \sum_{j=0}^{k} \binom{2n}{j} = \mathrm{O}(n^k)$ such terms, Step 1 can be done in $\mathrm{O}(n^{k+1}|\widetilde{F}|)$ time. In Step 2, we solve a k-TAUTOLOGY for each $a \in \widetilde{T}$ to check whether $\varphi_{A(a)} = \top$ holds. Hence if $k \leqslant 2$, this can be solved in $\mathrm{O}(|\varphi_{A(a)}|)$ time [4], where $|\varphi|$ denotes the number of literals in φ. Since $\varphi_{A(a)}$ can be constructed from φ in $\mathrm{O}(|\varphi|)$ time and $|\varphi_{A(a)}| \leqslant |\varphi| = \mathrm{O}(kn^k)$ holds, Step 2 can be done in $\mathrm{O}(kn^k|\widetilde{T}|)$ time. Totally, CHECK-RE($\mathcal{C}_{k\text{-DNF}}$) can be executed in $\mathrm{O}(n^k(k|\widetilde{T}| + n|\widetilde{F}|))$ time. □

For $k \geqslant 3$, however, CHECK-RE($\mathcal{C}_{k\text{-DNF}}$) does not run in polynomial time since it must check if $\varphi_{A(v)} = \top$, which is co-NP-complete. In fact, we have the next theorem.

Theorem 15. *For any fixed $k \geqslant 3$, problem* RE($\mathcal{C}_{k\text{-DNF}}$) *is* co-NP-*complete.*

Proof. Apply algorithm CHECK-RE($\mathcal{C}_{k\text{-}DNF}$) given in the proof of Theorem 14. Step 1 is carried out in polynomial time as noted therein. Step 2 consists of checking if $\varphi_{A(a)} = \top$ for polynomially many a, each of which is obviously a computation in co-NP. Therefore, RE($\mathcal{C}_{k\text{-DNF}}$) for $k \geqslant 3$ belongs to co-NP. The proof for the co-NP-hardness can be done by a reduction from the non-2-colorability problem of 3-uniform hypergraphs [21] (see [12] for details). □

We now turn to problem CE($\mathcal{C}_{k\text{-DNF}}$) for a fixed k, and first consider problem CE($\mathcal{C}_{1\text{-DNF}}$). Let $(\widetilde{T}, \widetilde{F})$ be a pBmb, where $\widetilde{T}, \widetilde{F} \subseteq \mathbb{M}^V$ and $V = \{1, 2, \ldots, n\}$. For a vector $v \in \mathbb{M}^V$ and a subset $I \subseteq V$, let $v[I]$ denote the projection of v on I. Furthermore, for a set $\widetilde{S} \subseteq \mathbb{M}^V$ and a subset $I \subseteq V$, let $\widetilde{S}[I]$ denote the projection of $\widetilde{S}$ on I (we assume that this projection keeps its multiplicity). If I is a singleton, say $I = \{j\}$, we write simply $\widetilde{S}[j]$ instead of $\widetilde{S}[\{j\}]$.

We shall show that the following algorithm can solve problem CE($\mathcal{C}_{1\text{-DNF}}$) in polynomial time.

Algorithm FIND-CE($\mathcal{C}_{1\text{-DNF}}$)

Input: A pBmb $(\widetilde{T}, \widetilde{F})$, where $\widetilde{T}, \widetilde{F} \subseteq \mathbb{M}^V$ and $V = \{1, 2, \ldots, n\}$.

Output: If the pBmb $(\widetilde{T}, \widetilde{F})$ has a consistent extension in $\mathcal{C}_{1\text{-DNF}}$, then output an assignment $\alpha \in \mathbb{B}^{AS}$ such that $(\widetilde{T}^\alpha, \widetilde{F}^\alpha)$ has an extension in $\mathcal{C}_{1\text{-DNF}}$ together with its 1-DNF expression φ; otherwise, NO.

Step 1. Let $I_0 := \{j \in V \mid 0 \in \widetilde{F}[j], 1 \notin \widetilde{F}[j]\}$, $I_1 := \{j \in V \mid 1 \in \widetilde{F}[j], 0 \notin \widetilde{F}[j]\}$, $I_{01} := \{j \in V \mid 0, 1 \in \widetilde{F}[j]\}$, and $I := V \setminus (I_0 \cup I_1 \cup I_{01})$ (i.e., $\widetilde{F}[j]$ for $j \in I$ contains only $*$). Define an assignment α by

$$\alpha(a, j) := \begin{cases} 1 & \text{if either (i) } j \in I_{01}, \text{ or (ii) } a \in \widetilde{T} \text{ and } j \in I_0, \text{ or} \\ & \quad \text{(iii) } a \in \widetilde{F} \text{ and } j \in I_1, \\ 0 & \text{if either (iv) } a \in \widetilde{F} \text{ and } j \in I_0, \text{ or (v) } a \in \widetilde{T} \text{ and } j \in I_1, \end{cases} \tag{24}$$

and 1-DNF

$$\varphi := \bigvee_{i \in I_0} x_i \vee \bigvee_{i \in I_1} \bar{x}_i. \tag{25}$$

Step 2. Define a pBmb $(\widetilde{T}', \widetilde{F}')$ with $\widetilde{T}', \widetilde{F}' \subseteq \mathbb{M}^I$ by

$$\widetilde{T}' := (\widetilde{T} \setminus \widetilde{S}_1)[I], \qquad \widetilde{F}' := \widetilde{F}[I],$$

where I was defined in Step 1, and $\widetilde{S}_1 = \{a \in \widetilde{T} \mid a_j \in \{1, *\}$ for some $j \in I_0\} \cup \{a \in \widetilde{T} \mid a_j \in \{0, *\}$ for some $j \in I_1\}$.

Step 3. For each $j \in I$, introduce a binary variable y_j (these variables define an assignment $\beta \in \mathbb{B}^{AS(\widetilde{T}' \cup \widetilde{F}')}$ such that $\beta(a, j) = y_j$ for all $(a, j) \in AS(\widetilde{T}')$ and $\beta(b, j) = \bar{y}_j$ for all $(b, j) \in AS(\widetilde{F}')$). Let $\widetilde{T}'' := \widetilde{T}' \cap \mathbb{B}^I$, and construct a CNF (conjunctive normal form)

$$\Phi(y) = \bigwedge_{a \in \widetilde{T}''} C_a, \quad C_a = \bigvee_{j \in ON(a)} y_j \vee \bigvee_{j \in OFF(a)} \bar{y}_j.$$

Find a solution satisfying $\Phi(y) = 1$ (i.e., solve problem SAT). If there exists a solution y^*, then let $\varphi' = \bigvee_{j \in ON(y^*)} x_j \vee \bigvee_{j \in OFF(y^*)} \bar{x}_j$, and output $\varphi := \varphi \vee \varphi'$ and the concatenated assignment (α, β), where β is obtained by substituting $y_j = y_j^*$ in the way as shown above; otherwise, output NO.

To see the correctness of algorithm FIND-CE($\mathcal{C}_{1\text{-DNF}}$), let us show the following lemma.

Lemma 6. *A pBmb $(\widetilde{T}, \widetilde{F})$ has a consistent extension in $\mathcal{C}_{1\text{-DNF}}$ if and only if $(\widetilde{T}', \widetilde{F}')$ obtained in Step 2 of* FIND-CE($\mathcal{C}_{1\text{-DNF}}$) *has a consistent extension in $\mathcal{C}_{1\text{-DNF}}$.*

Proof. Let φ be the 1-DNF of (25), and let φ' be a 1-DNF consistent extension of $(\widetilde{T}', \widetilde{F}')$. Then we claim that the 1-DNF $\varphi \vee \varphi'$ defines a consistent extension of $(\widetilde{T}, \widetilde{F})$, which will prove the if part. By the assignment α of (24), $\varphi(a^\alpha) = 1$ holds for all $a \in \widetilde{S}_1$, and $\varphi(b^\alpha) = 0$ holds for all $b \in \widetilde{F}$. Furthermore, since φ' is a consistent extension of $(\widetilde{T}', \widetilde{F}')$, some assignment $\beta \in \mathbb{B}^{AS(\widetilde{T}' \cup \widetilde{F}')}$ satisfies $\varphi(a^\beta) = 1$ for all $a \in \widetilde{T}'$ and $\varphi(b^\beta) = 0$ for all $b \in \widetilde{F}'$. Hence, by the definition of $\widetilde{F}'$, $\varphi(b^{(\alpha,\beta)}) = 0$ holds for all $b \in \widetilde{F}$. This implies that $\varphi \vee \varphi'$ is a 1-DNF extension of $(\widetilde{T}^{(\alpha,\beta)}, \widetilde{F}^{(\alpha,\beta)})$, that is, $\varphi \vee \varphi'$ is a 1-DNF consistent extension of $(\widetilde{T}, \widetilde{F})$.

Conversely, let $\gamma \in \mathbb{B}^{AS}$ be an assignment such that $(\widetilde{T}^\gamma, \widetilde{F}^\gamma)$ has a 1-DNF extension

$$\varphi^* = \bigvee_{i \in P} x_i \vee \bigvee_{i \in N} \bar{x}_i.$$

Then the following properties hold:

$$I_{01} \cap (P \cup N) = \emptyset, \qquad I_0 \cap N = \emptyset, \qquad I_1 \cap P = \emptyset,$$

since otherwise some vector $b \in \widetilde{F}$ would satisfy $f(b^\gamma) = 1$, a contradiction. Let

$$\varphi' = \bigvee_{i \in P \setminus I_0} x_i \vee \bigvee_{i \in N \setminus I_1} \bar{x}_i,$$

and let $\beta = \gamma[AS(\widetilde{T}' \cup \widetilde{F}')]$ (i.e., $\beta \in \mathbb{B}^{AS(\widetilde{T}' \cup \widetilde{F}')}$ is the projection of γ on $AS(\widetilde{T}' \cup \widetilde{F}')$). By the above properties, φ' is defined on I. We now show that φ' is an extension of $((\widetilde{T}')^\beta, (\widetilde{F}')^\beta)$, which will prove the only-if part. By the definition of φ', all $b \in \widetilde{F}'$ satisfy $\varphi'(b^\beta) = 0$. Assume that $a[I] \in \widetilde{T}'$ of some $a \in \widetilde{T}$ satisfies $\varphi'(a^\beta) = 0$. Then $(\bigvee_{i \in I_0} x_i \vee \bigvee_{i \in I_1} \bar{x}_i)(a^\beta) = 1$ holds by $\varphi^*(a^\gamma) = 1$ and by the definition of φ'. However, since $\widetilde{T}' = (\widetilde{T} \setminus S_1)[I]$, we have $a_j = 0$ for $j \in I_0$ and $a_j = 1$ for $j \in I_1$, which is a contradiction. Hence φ' is an extension of $((\widetilde{T}')^\beta, (\widetilde{F}')^\beta)$. □

Let us now consider how to obtain a consistent extension of $(\widetilde{T}', \widetilde{F}')$, that is, an assignment $\beta \in \mathbb{B}^{AS(\widetilde{T}' \cup \widetilde{F}')}$ such that $\big((\widetilde{T}')^\beta, (\widetilde{F}')^\beta\big)$ has a 1-DNF extension. Note that $AS(b) = I$ holds for all $b \in \widetilde{F}'$, i.e., all vectors in $\widetilde{F}'$ are $\{(*, *, \ldots, *)\}$. Furthermore, if $\beta(b, j) = 1$ (respectively, 0) holds for some $b \in \widetilde{F}'$, then no 1-DNF extension φ' of $((\widetilde{T}')^\beta, (\widetilde{F}')^\beta)$ has term x_j (respectively, $\bar{x}_j$). Since $\varphi'(a^\beta) = 1$ must holds for all $a \in \widetilde{T}'$, we would like to make $|T(\varphi')|$ as large as possible, under the condition that $\varphi'(b^\beta) = 0$ holds for all $b \in \widetilde{F}'$. This means that we only need to consider an assignment $\beta \in \mathbb{B}^{AS(\widetilde{T}' \cup \widetilde{F}')}$ such that $\beta(a, j) = y_j$ for all $(a, j) \in AS(\widetilde{T}')$ and $\beta(b, j) = \bar{y}_j$ for all $(b, j) \in AS(\widetilde{F}')$, where $y \in \mathbb{B}^I$, and a 1-DNF

$$\varphi' = \bigvee_{j \in ON(y)} x_j \vee \bigvee_{j \in OFF(y)} \bar{x}_j \tag{26}$$

is an extension of $((\widetilde{T}')^\beta, (\widetilde{F}')^\beta)$. For φ' to be an extension, we must choose a $y \in \mathbb{B}^I$ such that $\varphi'(a^\beta) = 1$ for all $a \in \widetilde{T}' \cap \mathbb{B}^I$. This condition can be written as $\Phi(y) = 1$ in Step 3. Therefore, $(\widetilde{T}', \widetilde{F}')$ has a 1-DNF consistent extension φ' of (26) if and only if $\Phi(y) = 1$ holds.

Theorem 16. *Problem* CE($\mathcal{C}_{1\text{-DNF}}$) *can be solved in polynomial time.*

Proof. The correctness of algorithm FIND-CE($\mathcal{C}_{1\text{-DNF}}$) is immediate from the above discussion. Therefore, let us consider its time complexity. Obviously, we can execute Steps 1 and 2 in $O(n(|\widetilde{T}| + |\widetilde{F}|))$ time. In Step 3, we must find a solution of $\Phi(y) = \bigwedge_{a \in \widetilde{T}''} C_a = 1$ (i.e., by solving an exact $|I|$-SAT, where exact k-SAT is a SAT satisfying that each of clauses has exact k literals). Exact k-SAT is in general NP-complete, but in this case, $k = |I|$, that is, k is equal to the dimension of SAT. Hence, this can be solved by checking if the number of different vectors in $\widetilde{T}''$ is equal to $2^{|I|}$ (in this case, $\Phi(y)$ is not satisfiable). Otherwise, we find a vector $y^* \in \mathbb{B}^I$ such that $\overline{y^*} \notin \widetilde{T}''$, and this y^* is a solution to $\Phi(y) = 1$. This can be done in $O(n|\widetilde{T}''|)$ time by using the binary tree of [33] as a data structure. In total, we need $O(n(|\widetilde{T}| + |\widetilde{F}|))$ time. □

For $k \geqslant 2$, however, we have the following negative result.

Theorem 17. *For any fixed $k \geqslant 2$, problem* $\mathrm{CE}(\mathcal{C}_{k\text{-DNF}})$ *is* NP*-complete, even if* $|AS(a)| \leqslant 1$ *holds for all* $a \in \widetilde{T} \cup \widetilde{F}$.

Proof. This problem is obviously in NP. To show its NP-hardness, let

$$\Phi = \bigwedge_{i=1}^{m} C_i, \quad C_i = (u_i \vee v_i \vee w_i), \ i = 1, 2, \ldots, m$$

be a cubic CNF, where u_i, v_i and w_i for $i = 1, 2, \ldots, m$ are literals from the set $L = \{x_1, \bar{x}_1, \ldots, x_n, \bar{x}_n\}$. We use notation like $x_j \in C_i$ by regarding C_i as the set $\{u_i, v_i, w_i\}$. Let $V_1 = \{1, 2, \ldots, n\}$, $V_2 = \{n+1, n+2, \ldots, n+m\}$, $V_3 = \{n+m+1, n+m+2, \ldots, n+m+k-2\}$ and $V = V_1 \cup V_2 \cup V_3$. We construct $\widetilde{T}, \widetilde{F} \subseteq \mathbb{M}^V$ as follows.

$$\begin{aligned}
\widetilde{T} &= \bigl\{a^{(i)} = \bigl(W_i \cup \{n+i\}; \emptyset\bigr) \mid i = 1, 2, \ldots, m\bigr\}, \\
\widetilde{F} &= \bigl\{(\emptyset; \emptyset), (V_1; \emptyset)\bigr\} \cup \bigl\{(\{j\}; \emptyset) \mid j \in V_1\bigr\} \\
&\qquad \cup \bigl\{b^{(i)} = \bigl(W_i \cup \{n+i\} \cup \{l\}; \emptyset\bigr) \mid i = 1, 2, \ldots, m, \ l \in V_3\bigr\} \\
&\qquad \cup \bigl\{c^{(i)} = \bigl(W_{\bar{i}} \cup \{n+i\}; \emptyset\bigr) \mid i = 1, 2, \ldots, m\bigr\} \\
&\qquad \cup \bigl\{d^{(j)} = \bigl(U_j; \{j\}\bigr) \mid j \in V_1\bigr\},
\end{aligned}$$

where $W_i = \{j \mid x_j \in C_i\}$, $W_{\bar{i}} = \{j \mid \bar{x}_j \in C_i\}$ and $U_j = \{n+i \mid x_j \in C_i \text{ or } \bar{x}_j \in C_i\}$. As before, $(R; S)$ denotes the vector $v \in \mathbb{M}^V$ such that $ON(v) = R$ and $AS(v) = \{(v, j) \mid j \in S\}$. Note that V_3 and the set of $b^{(i)}$ in $\widetilde{F}$ are both empty if $k = 2$. It is easy to see that $|AS(a)| \leqslant 1$ holds for all $a \in \widetilde{T} \cup \widetilde{F}$. We claim that this pBmb $(\widetilde{T}, \widetilde{F})$ has a consistent k-DNF extension if and only if 3-SAT for Φ has a solution (i.e., if there is a binary vector $y \in \{0, 1\}^n$ for which $\Phi(y) = 1$). This will complete the proof, because 3-SAT is NP-complete [21].

To prove the claim, let $\alpha \in \mathbb{B}^{AS}$ be an assignment such that $(\widetilde{T}^{\alpha}, \widetilde{F}^{\alpha})$ has a k-DNF extension φ, and let $t_i = \prod_{j \in P_i} x_j \prod_{j \in N_i} \bar{x}_j$, where $P_i \cap N_i = \emptyset$ and $|P_i \cup N_i| \leqslant k$, be a term in φ such that $t_i(a^{(i)}) = 1$ for $a^{(i)} \in \widetilde{T}$. Then such terms t_i, $i = 1, 2, \ldots, m$, satisfy the following properties:

(a) $N_i \supseteq V_3$ and $|(P_i \cup N_i) \setminus V_3| \leqslant 2$ hold. If the first property does not hold, we have $t_i(b^{(i)}) = 1$, a contradiction. The second property then follows from $|V_3| = k - 2$.

(b) $|(P_i \cup N_i) \cap V_2| = 1$ holds. Otherwise, we have $|(P_i \cup N_i) \cap V_2| = 0$ or 2 by (a). If $|(P_i \cup N_i) \cap V_2| = 0$, then at least one vector b in $\{(\emptyset; \emptyset), (V_1; \emptyset)\} \cup \{(\{j\}; \emptyset) \mid j \in V_1\}$ ($\subseteq \widetilde{F}$) satisfies $t_i(b) = 1$, which is a contradiction. Furthermore, if $|(P_i \cup N_i) \cap V_2| = 2$, then $|(P_i \cap N_i) \cap V_1| = 0$ holds, and $t_i(a^{(i)}) = 1$ implies that $c^{(i)} \in \widetilde{F}$ satisfies $t_i(c^{(i)}) = 1$, which is again a contradiction.

(c) $P_i \cap V_2 = \{n+i\}$ holds. First if $|P_i \cap V_2| = 1$, then $t_i(a^{(i)}) = 1$ implies $P_i \cap V_2 = \{n+i\}$. Otherwise, (b) implies $N_i \cap V_2 = \{n+h\}$ for some $h \in \{1, 2, \ldots, m\}$ with $h \neq i$, and then, by $|(P_i \cup N_i) \cap V_1| \leqslant 1$, at least one vector of $b = (\emptyset; \emptyset)$ and $b = (V_1; \emptyset)$ ($\in \widetilde{F}$) satisfies $t_i(b) = 1$, which is a contradiction. This means that either $t_i = x_{n+i}(\prod_{l \in V_3} \bar{x}_l)$ or $t_i = z_j x_{n+i}(\prod_{l \in V_3} \bar{x}_l)$ with $z_j \in L$ holds.

(d) $t_i = u_i x_{n+i}(\prod_{l \in V_3} \bar{x}_l)$, $v_i x_{n+i}(\prod_{l \in V_3} \bar{x}_l)$ or $w_i x_{n+i}(\prod_{l \in V_3} \bar{x}_l)$ holds. If $t_i = x_{n+i}(\prod_{l \in V_3} \bar{x}_l)$, then $c^{(i)} \in \widetilde{F}$ satisfies $t_i(c^{(i)}) = 1$, which is a contradiction. Therefore

$$t_i = z_j x_{n+i}\left(\prod_{l \in V_3} \bar{x}_l\right)$$

for some $z_j \in L$. If $z_j \in L \setminus \{u_i, v_i, w_i\}$, then z_j must be a negative literal x_l in order to have $t_i(a^{(i)}) = 1$, and $l \notin W_{\bar{i}}$ holds by the assumption on z_j. This means $t_i(c^{(i)}) = 1$, which is again a contradiction.

(e) There is no pair of terms t_p and t_q such that $t_p = x_j x_{n+p}(\prod_{l \in V_3} \bar{x}_l)$ and $t_q = \bar{x}_j x_{n+q}(\prod_{l \in V_3} \bar{x}_l)$. If t_p and t_q are such terms, then $(t_p \vee t_q)((d^{(j)})^{\beta}) = 1$ holds for any assignment $\beta \in \mathbb{B}^{AS}$, which is a contradiction.

Based on these, let us define a binary vector $y \in \mathbb{B}^n$ by

$$y_j = \begin{cases} 1 & \text{if } t_i = x_j x_{n+i}(\prod_{l \in V_3} \bar{x}_l) \text{ for some } i \in \{1, 2, \ldots, m\}, \\ 0 & \text{otherwise.} \end{cases}$$

Then properties (d) and (e) show that this y satisfies $\Phi(y) = 1$.

To prove the converse direction, given a binary vector $y \in \mathbb{B}^n$ satisfying $\Phi(y) = 1$, define an assignment $\alpha \in \mathbb{B}^{AS}$ by

$$\alpha(d^{(j)}, j) = \begin{cases} 1 & \text{if } y_j = 0, \\ 0 & \text{otherwise,} \end{cases}$$

and a k-DNF function φ^* by

$$\varphi^* = \bigvee_{i=1}^{m} t_i^*, \quad t_i^* = z_j x_{n+i}\left(\prod_{l \in V_3} \bar{x}_l\right), \tag{27}$$

where $z_j \in \{u_i, v_i, w_i\} = C_i$ and $z_j = 1$ is implied by y. Then we can see that φ^* is an extension of $(\widetilde{T}^{\alpha}, \widetilde{F}^{\alpha})$, that is, φ^* is a consistent extension of $(\widetilde{T}, \widetilde{F})$. □

4. Problem MRE

In this section we study the problem of finding most robust extensions. Let us recall that whenever problem CE($\mathcal{C}$) is NP-complete, problem MRE($\mathcal{C}$) must also be NP-hard. This leaves unresolved only about half of the considered cases. We shall see in the next subsection that some other cases can also be resolved as direct consequences of BEST-FIT. The remaining cases will then be discussed in the subsequent subsections.

4.1. Implications by problem BEST-FIT

Theorem 18. *If $\mathcal{C} \subseteq \mathcal{C}^+$ and problem* BEST-FIT($\mathcal{C}$) *can be solved in polynomial time, then* MRE($\mathcal{C}$) *can be solved in polynomial time for any pBmb $(\widetilde{T}, \widetilde{F})$ that satisfies $|AS(v)| \leqslant 1$ for all $v \in \widetilde{T} \cup \widetilde{F}$.*

Proof. For the above pBmb $(\widetilde{T}, \widetilde{F})$, define the pdBf (T', F') by $T' = \{a^{\mathbf{0}}, a^{\mathbf{1}} \mid a \in \widetilde{T}\}$ and $F' = \{b^{\mathbf{0}}, b^{\mathbf{1}} \mid b \in \widetilde{F}\}$, where $c^{\mathbf{0}}$ and $c^{\mathbf{1}}$ are defined as in Lemma 1. Also define the weights of the vectors in T' and F' by

$$\begin{array}{ll} w(a^{\mathbf{1}}) = +\infty & \text{if } a \in \widetilde{T}, \\ w(b^{\mathbf{0}}) = +\infty & \text{if } b \in \widetilde{F}, \\ w(a^{\mathbf{0}}) = 1 & \text{if } a \in \widetilde{T} \text{ and } AS(a) \neq \emptyset, \\ w(b^{\mathbf{1}}) = 1 & \text{if } b \in \widetilde{F} \text{ and } AS(a) \neq \emptyset. \end{array}$$

Note that by definition, $w(a^{\mathbf{1}}) = w(a) = +\infty$ (respectively, $w(b^{\mathbf{0}}) = w(b) = +\infty$) holds for all vectors $a \in \widetilde{T}$ (respectively, $b \in \widetilde{F}$) having no missing bit. We claim that

$$\rho(\widetilde{T}, \widetilde{F}) = \varepsilon(T', F')$$

holds, where ρ and ε are defined in (4) and (3), respectively. This will prove the theorem since $\varepsilon(T', F')$ can be computed by solving BEST-FIT($\mathcal{C}$).

First, if $\varepsilon(T', F') < +\infty$, then clearly, there is a consistent extension of $(\widetilde{T}, \widetilde{F})$ by the definition of w. Conversely, if there is a consistent extension f of $(\widetilde{T}, \widetilde{F})$, then $f(a^{\mathbf{1}}) = 1$ holds for all $a \in \widetilde{T}$ and $f(b^{\mathbf{0}}) = 0$ holds all $b \in \widetilde{F}$ by the positivity of f, which implies $\varepsilon(T', F') < +\infty$.

Let us assume next that there is a solution of MRE($\mathcal{C}$) for $(\widetilde{T}, \widetilde{F})$; i.e., a subset $Q \subseteq AS$ with $|Q| = \rho(\widetilde{T}, \widetilde{F})$ and an assignment $\beta \in \mathbb{B}^Q$ for which $(\widetilde{T}^{\beta}, \widetilde{F}^{\beta})$ has a robust extension f in $\mathcal{C}$. Then f correctly classifies all vectors in $T' \cup F'$, except for $a^{\bar{\beta}} \in T' \cup F'$ with $AS(a) \cap Q \neq \emptyset$ (where $\bar{\beta}$ denotes the complement of β). Hence

$$\rho(\widetilde{T}, \widetilde{F}) = |Q| = \sum_{a^{\mathbf{0}} \in T' \text{ s.t. } f(a^{\mathbf{0}})=0} w(a^{\mathbf{0}}) + \sum_{b^{\mathbf{1}} \in F' \text{ s.t. } f(b^{\mathbf{1}})=1} w(b^{\mathbf{1}}) \geqslant \varepsilon(T', F').$$

For the converse inequality, consider a solution (T^*, F^*) to BEST-FIT($\mathcal{C}$) for the pdBf (T', F'), i.e., $T^* \cap F^* = \emptyset$, $T^* \cup F^* = T' \cup F'$, the pdBf (T^*, F^*) has an extension f in $\mathcal{C}$, and $\varepsilon(T', F') = w(T' \cap F^*) + w(F' \cap T^*) < +\infty$. Then, by the positivity of f, we have $a^{\mathbf{1}} \in T^*$ for all $a \in \widetilde{T}$ and $b^{\mathbf{0}} \in F^*$ for all $b \in \widetilde{F}$. Thus define $Q = Q_1 \cup Q_0$, where

$$Q_1 = \{(a, j) \mid a \in T', (a, j) \in AS(a) \text{ and } a^{\mathbf{0}} \in F^*\},$$
$$Q_0 = \{(b, j) \mid b \in F', (b, j) \in AS(b) \text{ and } b^{\mathbf{1}} \in T^*\},$$

and an assignment $\beta \in \mathbb{B}^Q$ by $\beta(a, j) = 1$ for $(a, j) \in Q_1$, and $\beta(b, j) = 0$ for $(b, j) \in Q_0$. The resulting $(\widetilde{T}^{\beta}, \widetilde{F}^{\beta})$ has a robust extension $f \in \mathcal{C}$. Consequently,

$$\varepsilon(T', F') = |Q_1| + |Q_0| = |Q| \geqslant \rho(\widetilde{T}, \widetilde{F}). \quad \square$$

Combining these with the results in [11], we obtain the next corollary.

Corollary 4. *Let a pBmb $(\widetilde{T}, \widetilde{F})$ satisfy $|AS(a)| \leqslant 1$ for all $a \in \widetilde{T} \cup \widetilde{F}$. Then problem* MRE($\mathcal{C}$) *is polynomially solvable for the class of positive functions $\mathcal{C} = \mathcal{C}^+$.*

4.2. General extensions

Let us consider the class $\mathcal{C}_{all}$. As a result of Theorem 5, problem MRE($\mathcal{C}_{\text{all}}$) is NP-hard, unless instances $(\widetilde{T}, \widetilde{F})$ satisfy

$$|AS(a)| \leqslant 1 \quad \text{for all } a \in \widetilde{T} \cup \widetilde{F}.$$

We shall show below that, for such pBmb instances, a most robust extension can be found in polynomial time.

Let us remark first that any assignment $\alpha \in \mathbb{B}^{AS}$ for which $(\widetilde{T}^{\alpha}, \widetilde{F}^{\alpha})$ has an extension must satisfy the conditions (i) and (ii) in the proof of Theorem 4. Hence, some components of such an α may be forced to take a unique binary value by conditions (8) and (9). Let us assume therefore that we fix all such missing $*$ bits in advance, and let us consider only conditions (10) and (11) in the sequel.

Let us define next a bipartite multi-graph with labeled edges.

$$\begin{aligned} &G_{AS} = (V, E), \\ &V = AS(\widetilde{T}) \cup AS(\widetilde{F}), \\ &E = \big\{(q, r; \alpha) \mid q = (a, i) \in AS(\widetilde{T}), r = (b, j) \in AS(\widetilde{F}), \text{ there exists} \\ &\qquad\qquad \text{an assignment } \alpha \in \mathbb{B}^{\{q,r\}} \text{ such that } a^{\alpha} = b^{\alpha}\big\}. \end{aligned} \tag{28}$$

The label $c(e)$ of each edge $e = (q, r; c(e))$, as defined in (28), is called the *configuration* of e. Let us note that, since $|AS(a)| \leqslant 1$ holds for all $a \in \widetilde{T} \cup \widetilde{F}$, every pair of $q = (a, i) \in AS(\widetilde{T})$ and $r = (b, j) \in AS(\widetilde{F})$ has at most two assignments $\alpha \in \mathbb{B}^{\{q,r\}}$ such that $a^{\alpha} = b^{\alpha}$. If there are two assignments $\alpha \in \mathbb{B}^{\{q,r\}}$ for some $q \in AS(\widetilde{T})$ and $r \in AS(\widetilde{F})$, for which $a^{\alpha} = b^{\alpha}$ (this occurs if $q = (a, i)$ and $r = (b, j)$ satisfy $i = j$), then the graph G_{AS} has parallel edges corresponding to such different configurations.

Example 3. Let us define $\widetilde{T}, \widetilde{F} \subseteq \{0, 1\}^6$ by

$$\widetilde{T} = \left\{ \begin{array}{l} a^{(1)} = (*, 1, 1, 1, 1) \\ a^{(2)} = (1, 1, 1, 1, *) \\ a^{(3)} = (1, 1, 1, *, 1) \\ a^{(4)} = (1, 1, *, 1, 1) \\ a^{(5)} = (1, *, 0, 1, 0) \end{array} \right\}, \qquad \widetilde{F} = \left\{ \begin{array}{l} b^{(1)} = (1, *, 1, 1, 1) \\ b^{(2)} = (1, 1, 1, 1, *) \\ b^{(3)} = (1, 1, *, 1, 0) \\ b^{(4)} = (1, 1, 0, 1, *) \end{array} \right\}.$$

Then the graph G_{AS} is given in Fig. 2. Although the configurations of edges are not indicated, they are easily found out. For example, the edge $e = ((a^{(1)}, 1), (b^{(1)}, 2))$ has $c(e) = (a_1^{(1)} = 1, b_2^{(1)} = 1)$, and the parallel edges $e' = ((a^{(2)}, 5), (b^{(2)}, 5))$ and $e'' = ((a^{(2)}, 5), (b^{(2)}, 5))$ have $c(e') = (a_5^{(2)} = 0, b_5^{(2)} = 0)$ and $c(e'') = (a_5^{(2)} = 1, b_5^{(2)} = 1)$, respectively.

Lemma 7. *Given a pBmb $(\widetilde{T}, \widetilde{F})$, an assignment $\beta \in \mathbb{B}^{Q}$ for a subset $Q \subseteq AS$ is a robust assignment of $(\widetilde{T}, \widetilde{F})$ (i.e., $(\widetilde{T}^{\beta}, \widetilde{F}^{\beta})$ has a robust extension) if and only if, for every edge*

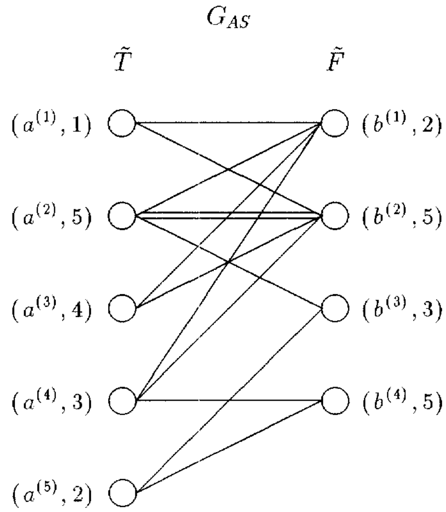

Fig. 2. The graph G_{AS} of the pBmb $(\tilde{T}, \tilde{F})$ in Example 3.

$e = (q, r; \alpha)$ of G_{AS}, we have either $q = (a, i) \in Q$ and $a^\beta \neq a^\alpha$, or $r = (b, j) \in Q$ and $b^\beta \neq b^\alpha$, or both.

Proof. Let us first show the only-if part. Let f be a robust extension of $(\tilde{T}^\beta, \tilde{F}^\beta)$, and let $e = (q, r; \alpha)$ be an edge of G_{AS}. Assume that either $q \notin Q$ or $a^\beta = a^\alpha$ holds. Then we have $f(a^\alpha) = 1$. Indeed, if $q = (a, i) \notin Q$, then $(a^\beta)^\alpha = a^\alpha$, and since $\beta \in \mathbb{B}^Q$ is a robust assignment, $f(a^\alpha) = 1$ must hold. On the other hand, if $a^\beta = a^\alpha$, then obviously $f(a^\alpha) = f(a^\beta) = 1$ must hold, since $a \in \tilde{T}$.

We then show that $f(a^\alpha) = 1$ implies $r = (b, j) \in Q$ and $b^\beta \neq b^\alpha$, which proves the only-if part. If $r \notin Q$, then $(b^\beta)^\alpha = b^\alpha = a^\alpha$, and hence $f(a^\alpha) = f(b^\alpha) = 0$ by $b \in \tilde{F}$, which is a contradiction. Similarly $b^\beta = b^\alpha$ leads to the same contradiction. Hence $r \in Q$ and $b^\beta \neq b^\alpha$ must hold.

To prove the if part, assume that $\beta \in \mathbb{B}^Q$ for a subset $Q \subseteq AS$ is not a robust assignment of $(\tilde{T}, \tilde{F})$. Then, by the definition of robustness, we have a pair of vectors $a \in \tilde{T}$ and $b \in \tilde{F}$ such that $a^\beta \approx b^\beta$. Then the edge $e = (q, r; \alpha)$ with $q = (a, i)$ and $r = (b, j)$ does not satisfy the statement of the lemma. □

For a vector $d \in \mathbb{B}^n$, let $E(d)$ denote the set of edges $e = (q, r; \alpha) \in E$ such that $a^\alpha = b^\alpha = d$, where $q = (a, i)$ and $r = (b, j)$. Then $E = \bigcup_d E(d)$. Let us define a *coherent domain* $D(d)$ as the set of vertices incident to some edges of $E(d)$, and let D_0 denote the set of isolated vertices (i.e., incident to no edge $e \in E$). No vertex in D_0 belongs to a coherent domain. In the following discussion, we only consider nonempty coherent domains. Fig. 3 shows all nonempty coherent domains of the graph G_{AS} of $(\tilde{T}, \tilde{F})$ in Example 3.

Lemma 8. *Every coherent domain $D(d) \subseteq V$ of G_{AS} induces a complete bipartite subgraph of G_{AS}.*

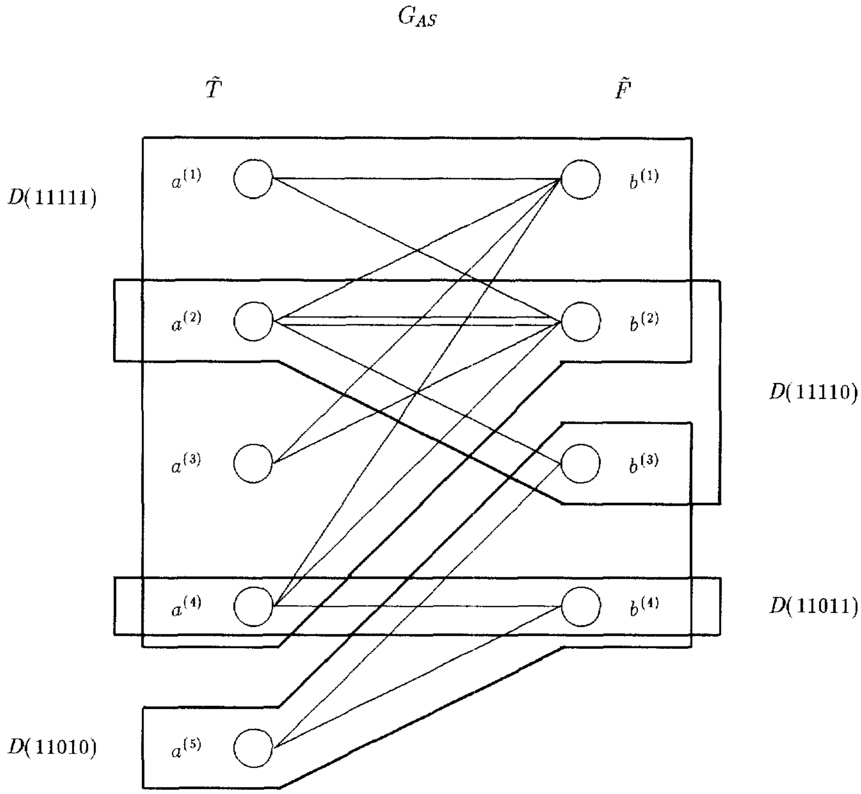

Fig. 3. Coherent domains of the graph G_{AS} of $(\widetilde{T}, \widetilde{F})$ in Example 3.

Proof. Take any pair $q = (a, i) \in AS(\widetilde{T})$ and $r = (b, j) \in AS(\widetilde{F})$ that satisfy $q, r \in D(d)$. Then there exist assignments $\alpha \in \mathbb{B}^{\{q\}}$ and $\beta \in \mathbb{B}^{\{r\}}$ such that $d = a^{\alpha} = b^{\beta}$. We concatenate these assignments to have an assignment $\gamma = (\alpha, \beta) \in \mathbb{B}^{\{q,r\}}$ for which $a^{\gamma} = b^{\gamma} = d$, implying that there is an edge $(q, r; \gamma) \in E(d)$. □

Lemma 9. *Let $D(d)$ and $D(d')$ be two coherent domains of G_{AS}, where $d, d' \in \mathbb{B}^n$ and $d \neq d'$. If $D(d) \cap D(d') \neq \emptyset$, then $\| d - d' \| = 1$ holds, where $\| x \| = \sum_{i=1}^{n} |x_i|$.*

Proof. Let $q = (a, i) \in D(d) \cap D(d')$. Then there exist two assignments $\alpha, \beta \in \mathbb{B}^{\{q\}}$ ($= \{0, 1\}$) such that $a^{\alpha} = d$ and $a^{\beta} = d'$. Since $|AS(a)| \leqslant 1$ is assumed, this implies $\|d - d'\| = 1$. □

Lemma 10. *Let $D(d)$ and $D(d')$ be two coherent domains of G_{AS}, where $d, d' \in \mathbb{B}^n$ and $d \neq d'$. Then $|D(d) \cap D(d')| \leqslant 2$ holds. Furthermore, if $D(d) \cap D(d') = \{q, r\}$, then the graph G_{AS} has two parallel edges between q and r.*

Proof. If $q = (a, i), r = (b, j) \in D(d) \cap D(d')$, then by assigning 0 and 1 to q and r, each of a and b can become both d and d'. Since $\| d - d' \| = 1$ by Lemma 9, this can only happen if the vectors a and b are identical, missing the same component $i = j$. Therefore $|D(d) \cap D(d') \cap AS(\widetilde{T})| \leqslant 1$ and $|D(d) \cap D(d') \cap AS(\widetilde{F})| \leqslant 1$, and hence $|D(d) \cap D(d')| \leqslant 2$. Finally, if $D(d) \cap D(d') = \{q, r\}$, where $q = (a, i) \in AS(\widetilde{T})$ and $r = (b, j) \in AS(\widetilde{F})$, then $q = r$ implies that there are two assignments $\alpha, \beta \in \mathbb{B}^{\{q,r\}}$ such that $a^\alpha = b^\alpha = d$ and $a^\beta = b^\beta = d'$, i.e., the graph G_{AS} has two parallel edges between q and r. □

Let us now color the edges of G_{AS} by "yellow" and "blue", so that all edges of a set $E(d)$ have the same color, and every pair of sets $E(d)$ and $E(d')$ with $D(d) \cap D(d') \neq \emptyset$ has different colors. We call such a two coloring *alternating*. The following lemma shows that an alternating coloring is always possible. Furthermore, it can be uniquely completed after fixing the color of one set $E(d)$ in each connected component of G_{AS}.

Lemma 11. *Let $D(d^{(0)}), D(d^{(1)}), \ldots, D(d^{(l)})$ denote a cycle of coherent domains such that $d^{(i-1)} \neq d^{(i)}$ and $D(d^{(i-1)}) \cap D(d^{(i)}) \neq \emptyset$ hold for all $i = 1, 2, \ldots, l-1$, and $D(d^{(l)}) = D(d^{(0)})$. Then l is even.*

Proof. Lemma 9 implies that $\| d^{(i-1)} - d^{(i)} \| = 1$ holds for all $i = 1, 2, \ldots, l-1$. Since $\| d^{(0)} - d^{(l)} \| = 0$ is even, l must be even. □

Finally, let us orient the edges of G_{AS} according to a given alternating coloring, as follows. Every yellow edge (q, r) is oriented from $q \in AS(\widetilde{T})$ to $r \in AS(\widetilde{F})$, and every blue edge (q, r) is oriented from $r \in AS(\widetilde{F})$ to $q \in AS(\widetilde{T})$. Let G'_{AS} denote the resulting directed graph. For example, Fig. 4 shows the directed graph G'_{AS} corresponding to the pBmb $(\widetilde{T}, \widetilde{F})$ of Example 3. Let us observe that every directed path of this graph is alternating in colors, and every alternating undirected path is either forward directed or backward directed.

The next lemma characterizes a robust assignment by a directed path of G'_{AS}.

Lemma 12. *Let $(\widetilde{T}, \widetilde{F})$ be a pBmb. Then $\beta \in \mathbb{B}^Q$ for $Q \subseteq AS$ is a robust assignment if and only if the following two properties hold for every directed path $q^{(0)} \xrightarrow[e_1]{} q^{(1)} \xrightarrow[e_2]{} q^{(2)} \rightarrow \cdots \rightarrow q^{(l-1)} \xrightarrow[e_l]{} q^{(l)}$ in G'_{AS}, where $q^{(i)} = (a^{(i)}, j_i)$ and $\alpha_i = c(e_i)$ for all i.*

(i) *If $q^{(0)} \notin Q$ or $(a^{(0)})^\beta = (a^{(0)})^{\alpha_1}$, then $q^{(i)} \in Q$ and $(a^{(i)})^\beta \neq (a^{(i)})^{\alpha_i}$ hold for all $i = 1, 2, \ldots, l$.*

(ii) *If $q^{(l)} \notin Q$ or $(a^{(l)})^\beta = (a^{(l)})^{\alpha_l}$ for l, then $q^{(i)} \in Q$ and $(a^{(i)})^\beta \neq (a^{(i)})^{\alpha_{i+1}}$ hold for all $i = 0, 1, \ldots, l-1$.*

Proof. We first prove the only-if part. For condition (i), we first consider $e_1 = (q^{(0)}, q^{(1)})$. By Lemma 7, $q^{(0)} \notin Q$ or $(a^{(0)})^\beta = (a^{(0)})^{\alpha_1}$ implies that $q^{(1)} \in Q$ and $(a^{(1)})^\beta \neq (a^{(1)})^{\alpha_1}$. Now, since $e_1 = (q^{(0)}, q^{(1)}) \in E(d)$ and $e_2 = (q^{(1)}, q^{(2)}) \in E(d')$ have different colors, we must have $d \neq d'$ and $q^{(1)} \in D(d) \cap D(d')$, and hence $\| d - d' \| = 1$ by Lemma 9. Therefore, $(a^{(1)})^\beta \neq (a^{(1)})^{\alpha_1} (= d)$ implies $(a^{(1)})^\beta = (a^{(2)})^{\alpha_2} (= d')$, and hence $q^{(2)}$ satisfies $q^{(2)} \in Q$ and $(a^{(2)})^\beta \neq (a^{(2)})^{\alpha_2}$ by Lemma 7. This assignment can proceed in a similar manner to $q^{(i)}$, $i = 2, 3, \ldots, l$. Case (ii) is similar to (i).

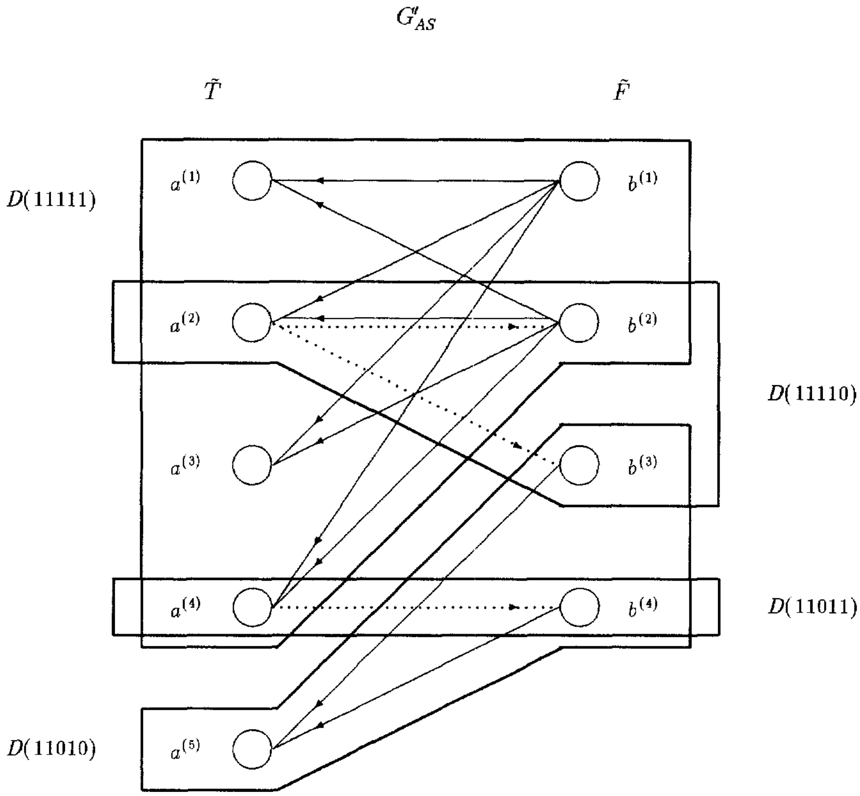

Fig. 4. The directed graph G'_{AS} of $(\widetilde{T}, \widetilde{F})$ in Example 3.

Conversely, if conditions (i) and (ii) hold, then, by Lemma 7, $\beta \in \mathbb{B}^Q$ is a robust assignment. □

Let C_i, $i = 1, 2, \ldots, s$, denote all the strongly connected components of this directed graph G'_{AS}. Furthermore, let G^*_{AS} denote the transitive closure of G'_{AS} (i.e., (s, t) is an arc in G^*_{AS} if there is an s–t directed path in G'_{AS}), and let G^0_{AS} denote the directed subgraph of G^*_{AS} induced by

$$W = \bigcup_{i \text{ s.t. } |C_i|=1} C_i. \tag{29}$$

It is easy to see that the set of isolated vertices D_0 in G_{AS} satisfies $D_0 \subseteq W$. Fig. 5 gives the graph G^0_{AS} of $(\widetilde{T}, \widetilde{F})$ in Example 3, in which arcs (u, v), having a directed path of length at least 2 from u to v, are not indicated for simplicity.

Lemma 13. *Let $(\widetilde{T}, \widetilde{F})$ be a pBmb, let $\alpha \in \mathbb{B}^Q$ for some $Q \subseteq AS$ be a robust assignment, and let C_i, W and G^0_{AS} be defined as above. Then the following two conditions hold:*

(i) *$C_i \subseteq Q$ for all C_i with $|C_i| > 1$, and*

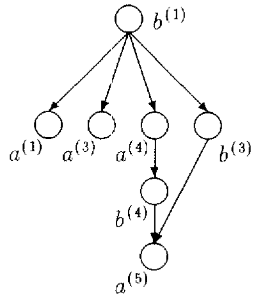

Fig. 5. The graph G^0_{AS} corresponding to G'_{AS} of $(\widetilde{T}, \widetilde{F})$ in Example 3.

(ii) *$W \setminus Q$ is an antichain in G^0_{AS} (i.e., for any pair of $q, r \in W \setminus Q$, there is no directed path from q to r in G_0, and vice versa).*

Proof. Consider a robust assignment $\alpha \in \mathbb{B}^Q$. Assume $q \in C_i \setminus Q$ for some C_i with $|C_i| > 1$. Then there is a directed cycle $q^{(0)}$ $(= q), q^{(1)}, q^{(2)}, \ldots, q^{(l)} (= q)$ of length $l > 1$ in G'_{AS}, and $q \notin Q$ implies $q \in Q$ by Lemma 12, which is a contradiction. Hence condition (i) holds. To prove condition (ii), let us assume that for some pair of $q, r \in W \setminus Q$, there exists a directed path from q to r in G'_{AS}. This is again a contradiction since $q \notin Q$ implies $r \in Q$ by Lemma 12. □

Lemma 14. *Let $(\widetilde{T}, \widetilde{F})$ be a pBmb, and let $S \subseteq W$ be any maximal antichain in G^0_{AS}. Then for $Q = AS \setminus S$, there is a robust assignment $\beta \in \mathbb{B}^Q$ of $(\widetilde{T}, \widetilde{F})$.*

Proof. For $Q = AS \setminus S$, we shall construct a robust assignment $\beta \in \mathbb{B}^Q$. In the following, we shall consider the directed graph G'_{AS}, and let us note that, by definition, S is also an antichain in G'_{AS}. Lemma 12 implies that, starting from a vertex $q \in S$ (i.e., $q \notin Q$), a robust assignment β for all vertices t which are either reachable from q or reachable to q is uniquely determined, unless the following cases of conflicts are encountered.

(i) For $q, r \in S$, there is a vertex t for which there are two directed paths $P_1 = q^{(0)}$ $(= q) \to q^{(1)} \to \cdots \to q^{(k)} (= t)$ and $P_2 = r^{(0)} (= r) \to r^{(1)} \to \cdots \to r^{(l)} (= t)$ such that $t^\alpha \neq t^{\alpha'}$, where $\alpha = c(q^{(k-1)}, t)$ and $\alpha' = c(r^{(l-1)}, t)$.

(ii) For $q, r \in S$, there is a vertex t for which there are two directed paths $P_1 = q^{(0)}$ $(= t) \to q^{(1)} \to \cdots \to q^{(k)} (= q)$ and $P_2 = r^{(0)} (= t) \to r^{(1)} \to \cdots \to r^{(l)} (= r)$ such that $t^\alpha \neq t^{\alpha'}$, where $\alpha = c(t, q^{(1)})$ and $\alpha' = c(t, r^{(1)})$.

If one of these conflicts occurs, Lemma 12 implies that t must be assigned in different ways, and hence we cannot construct an appropriate robust assignment β.

However, we now show that none of these conflicts can occur. Let us consider case (i) only, since case (ii) can be analogously treated. Now $t^\alpha \neq t^{\alpha'}$ implies $(q^{(k-1)}, t) \in E(d)$ and $(r^{(l-1)}, t) \in E(d')$ for some $d \neq d'$, and hence $t \in D(d) \cap D(d')$. Thus $(q^{(k-1)}, t)$ and $(r^{(l-1)}, t)$ have different colors, by $D(d) \cap D(d') \neq \emptyset$. By the rule of orienting edges (yellow edges are oriented from $AS(\widetilde{T})$ to $AS(\widetilde{F})$, and blue edges are oriented from $AS(\widetilde{F})$ to $AS(\widetilde{T})$), this means that one of $(q^{(k-1)}, t)$ and $(r^{(l-1)}, t)$ is oriented towards t, and the other is away from t, a contradiction to the assumption in (i).

Let us denote by R the set of all vertices $t \notin S$ such that either t is reachable from some $q \in S$ or some $q \in S$ is reachable from t. The above argument shows that a robust assignment β for R can be uniquely determined by Lemma 12. Finally, we consider an assignment $\beta \in \mathbb{B}^{AS\setminus(S\cup R)}$. By the maximality of S, every vertex $t \in AS \setminus (S \cup R)$ has an incoming arc $e = (r, t) \in E(d)$. Therefore, determine the robust assignment β of this t so that $t^{\beta} = d$ holds. This is well-defined because all incoming arcs to t belong to the same $E(d)$ by the definition of G'_{AS}. It is then easy by Lemma 7 to see that the resulting β over AS is in fact a robust assignment. $\square$

Lemmas 13 and 14 imply that problem MRE($\mathcal{C}_{\text{all}}$) is equivalent to the problem of finding a maximum antichain of G^0_{AS}. Since G^0_{AS} is acyclic, we can find such an antichain in polynomial time by applying a maximum flow algorithm (see Dilworth's theorem, e.g., [20]). Hence, we have shown the following theorem.

Theorem 19. *Problem* MRE($\mathcal{C}_{\text{all}}$) *can be solved in polynomial time for a pBmb* $(\widetilde{T}, \widetilde{F})$ *in which all* $a \in \widetilde{T} \cup \widetilde{F}$ *satisfy* $|AS(a)| \leqslant 1$.

4.3. Positive extensions

Let us consider the class $\mathcal{C}^+$ of positive functions. Corollary 4 implies us that problem MRE($\mathcal{C}^+$) can be solved in polynomial time for the restricted case of $|AS(a)| \leqslant 1$ for all $a \in \widetilde{T} \cup \widetilde{F}$. However, problem MRE($\mathcal{C}^+$) is in general NP-hard.

Theorem 20. *Problem* MRE($\mathcal{C}^+$) *is* NP*-hard, even if* $|AS(a)| \leqslant 2$ *holds for all* $a \in \widetilde{T} \cup \widetilde{F}$.

Proof. Let $G = (V, E)$ be an arbitrary graph, where $V = \{1, 2, \ldots, n\}$, and define $\widetilde{T}, \widetilde{F} \subseteq \mathbb{M}^V$ as follows.

$$\widetilde{T} = \{a^{(i,j)} = (\emptyset; \{i, j\}) \mid (i, j) \in E\},$$
$$\widetilde{F} = \{b^{(0)} = (\emptyset; \emptyset)\} \cup \{b^{(i)} = (\emptyset; \{i\}) \mid i \in V\},$$

where $(R; S)$ denotes the vector $v \in \mathbb{M}^V$ such that $ON(v) = R$ and $AS(v) = \{(v, j) \mid j \in S\}$. It is easy to see that $|AS(a)| \leqslant 2$ holds for all $a \in \widetilde{T} \cup \widetilde{F}$. We claim that

$$\rho(\widetilde{T}, \widetilde{F}) = |E| + \tau(G)$$

holds, where $\rho(\widetilde{T}, \widetilde{F})$ is defined by (4), and $\tau(G)$ denotes the cardinality of a minimum vertex cover of G. This will complete the proof of the theorem, since finding $\tau(G)$ is known to be NP-hard [21].

Let us first observe that, if $(\widetilde{T}^{\beta}, \widetilde{F}^{\beta})$ has a robust positive extension for some $\beta \in \mathbb{B}^Q$, $Q \subseteq AS$, then either $\beta(a^{(i,j)}, i) = 1$ or $\beta(a^{(i,j)}, j) = 1$ (or both) holds for every $(i, j) \in E$, since otherwise we have $b^{(0)} = (a^{(i,j)})^{\beta} \in \widetilde{F}$, which is a contradiction. Let

$$E_1 = \{(i, j) \in E \mid \text{exactly one of } \beta(a^{(i,j)}, i) = 1 \text{ and } \beta(a^{(i,j)}, j) = 1 \text{ holds}\},$$
$$E_2 = \{(i, j) \in E \mid \beta(a^{(i,j)}, i) = \beta(a^{(i,j)}, j) = 1\}.$$

If $(i,j) \in E_1$ and $\beta(a^{(i,j)}, i) = 1$ (respectively, $\beta(a^{(i,j)}, j) = 1$), then $\beta(b^{(i)}, i) = 0$ (respectively, $\beta(b^{(i)}, j) = 0$) (otherwise $(a^{(i,j)})^\beta \approx (b^{(i)})^\beta$ (respectively, $(a^{(i,j)})^\beta \approx (b^{(j)})^\beta$) and β is not a robust assignment). This implies that $C = \{i \mid \beta(b^{(i)}, i) = 0\} \cup \{i \mid i < j, (i,j) \in E_2\}$ is a vertex cover of G. Hence

$$\begin{aligned} |Q| &\geqslant |E_1| + 2|E_2| + \bigl|\{i \mid \beta(b^{(i)}, i) = 0\}\bigr| \\ &= \bigl(|E_1| + |E_2|\bigr) + \bigl(|E_2| + \bigl|\{i \mid \beta(b^{(i)}, i) = 0\}\bigr|\bigr) \\ &= |E| + |C| \geqslant |E| + \tau(G). \end{aligned}$$

For the converse direction, let $C \subseteq V$ be a minimum vertex cover, and let us define a set $Q \subseteq AS$ and an assignment $\beta \in \mathbb{B}^Q$ by

$$Q = \bigl\{(a^{(i,j)}, i) \mid \text{either } (i \in C,\ j \notin C) \text{ or } (i, j \in C, i < j)\bigr\} \cup \{(b^{(i)}, i) \mid i \in C\},$$

$\beta(a^{(i,j)}, i) = 1$ for $(a^{(i,j)}, i) \in Q$ and $\beta(b^{(i)}, i) = 0$ for $(b^{(i)}, i) \in Q$. By Lemma 1, it is easy to see that β is a robust assignment, and $|Q| = |E| + \tau(G)$ holds. □

4.4. Decomposable extensions

As in Section 3.5, we consider two basic classes of decomposable functions, $\mathcal{C}_{G(S_0, G(S_1))}$ and $\mathcal{C}^+_{G(S_0, G(S_1))}$. As noted at the end of Section 2.2, Theorem 13 immediately implies that problem MRE($\mathcal{C}_{G(S_0, G(S_1))}$) is NP-hard even if $|AS(a)| \leqslant 1$ holds for all $a \in \widetilde{T} \cup \widetilde{F}$. For problem MRE($\mathcal{C}^+_{G(S_0, G(S_1))}$), we also have the following negative result.

Theorem 21. *Problem* MRE($\mathcal{C}^+_{G(S_0, G(S_1))}$) *is* NP-*hard, even if* $|AS(a)| \leqslant 1$ *holds for all* $a \in \widetilde{T} \cup \widetilde{F}$.

Proof. The proof is done by modifying the NP-hardness proof of BEST-FIT($\mathcal{C}^+_{G(S_0, G(S_1))}$) in [11]. A complete proof can be found in [12]. □

4.5. k-DNF extensions

Let us consider the classes $\mathcal{C}_{k\text{-DNF}}$ and $\mathcal{C}^+_{k\text{-DNF}}$. For general k, these problems are NP-hard since the corresponding CEs are NP-complete by Corollary 1. Therefore, we consider only the cases of fixed k.

Theorem 22. *For any fixed* $k \geqslant 1$, *problems* MRE($\mathcal{C}_{k\text{-DNF}}$) *and* MRE($\mathcal{C}^+_{k\text{-DNF}}$) *are* NP-*hard, even if* $|AS(a)| \leqslant 1$ *holds for all* $a \in \widetilde{T} \cup \widetilde{F}$.

Proof. Let $G = (V, E)$ be a graph, where $V = \{1, 2, \ldots, n\}$, and let $W = \{n+1, n+2, \ldots, n+k-1\}$. Let us define $\widetilde{T}, \widetilde{F} \subseteq \mathbb{M}^{V \cup W}$ as follows.

$$\begin{aligned} \widetilde{T} &= \bigl\{a^{(i,j)} = \bigl(\{i,j\} \cup W; \emptyset\bigr) \mid (i,j) \in E\bigr\}, \\ \widetilde{F} &= \bigl\{b^{(0)} = (W; \emptyset)\bigr\} \cup \bigl\{b^{(i)} = (W; \{i\}) \mid i \in V\bigr\} \\ &\quad \cup \bigl\{b^{(i,j)} = \bigl(\{i,j\} \cup (W \setminus \{l\}); \emptyset\bigr) \mid (i,j) \in E,\ l \in W\bigr\}, \end{aligned}$$

where $(R; S)$ denotes the vector $v \in \mathbb{M}^{V \cup W}$ such that $ON(v) = R$ and $AS(v) = \{(v, j) \mid j \in S\}$. It is easy to see that $AS = \{(b^{(i)}, i) \mid i \in V\}$ and $|AS(a)| \leqslant 1$ for all $a \in \widetilde{T} \cup \widetilde{F}$ hold. We claim that

$$\rho(\mathcal{C}_{k\text{-DNF}}; (\widetilde{T}, \widetilde{F})) = \rho(\mathcal{C}^{+}_{k\text{-DNF}}; (\widetilde{T}, \widetilde{F})) = \tau(G) \tag{30}$$

holds, where $\tau(G)$ denotes the cardinality of a minimum vertex cover of graph G. This will complete the proof because finding $\tau(G)$ is known to be NP-hard [21].

To prove the claim, we show first that

$$\rho(\mathcal{C}_{k\text{-DNF}}; (\widetilde{T}, \widetilde{F})) \leqslant \rho(\mathcal{C}^{+}_{k\text{-DNF}}; (\widetilde{T}, \widetilde{F})) \leqslant \tau(G). \tag{31}$$

The first inequality follows from $\mathcal{C}_{k\text{-DNF}} \supseteq \mathcal{C}^{+}_{k\text{-DNF}}$. For the second one, let us associate a k-DNF φ_C to any subset $C \subseteq V$ by defining

$$\varphi_C = \bigvee_{i \in C} x_i x_{n+1} x_{n+2} \cdots x_{n+k-1},$$

and let us consider φ_{C^*}, where $C^* \subseteq V$ is a minimum vertex cover of G. Define $Q \subseteq AS$ and $\alpha \in \mathbb{B}^Q$ by $Q = \{(b^{(i)}, i) \mid i \in C^*\}$ and $\alpha((b^{(i)}, i)) = 0$ for all $(b^{(i)}, i) \in Q$, respectively. Then φ_{C^*} is a robust extension of $(\widetilde{T}^{\alpha}, \widetilde{F}^{\alpha})$, i.e., $\rho(\mathcal{C}^{+}_{k\text{-DNF}}; (\widetilde{T}, \widetilde{F})) \leqslant |C^*| = \tau(G)$.

Next, we show that

$$\rho(\mathcal{C}_{k\text{-DNF}}; (\widetilde{T}, \widetilde{F})) \geqslant \tau(G), \tag{32}$$

which together with (31) will imply (30). For this end, let $\alpha \in \mathbb{B}^Q$ for $Q \subseteq AS$ be an assignment such that $(\widetilde{T}^{\alpha}, \widetilde{F}^{\alpha})$ has a robust k-DNF extension, and let

$$\varphi = \bigvee_{i \in I} t_i,$$

be such a k-DNF with a minimal I, where $t_i = \prod_{j \in P_i} x_j \prod_{j \in N_i} \bar{x}_j$, $P_i \cap N_i = \emptyset$ and $|P_i \cup N_i| \leqslant k$ for all $i \in I$. Then the minimality of I implies that for every term t_i, there is an $a^{(h_i, l_i)} \in \widetilde{T}^{\alpha}$ such that $t_i(a^{(h_i, l_i)}) = 1$. Thus $P_i \supseteq W$ holds for every $i \in I$, since otherwise the vector $b^{(h_i, l_i)} \in \widetilde{F}^{\alpha}$ also satisfies $t_i(b^{(h_i, l_i)}) = 1$, which is a contradiction. This implies $|(P_i \cup N_i) \cap V| \leqslant 1$ by $|P_i \cup N_i| \leqslant k$. Furthermore, $|P_i \cap V| = 1$ holds for every $i \in I$; otherwise (i.e., $P_i \cap V = \emptyset$), $t_i(b^{(0)}) = 1$ holds for $b^{(0)} \in \widetilde{F}^{\alpha}$, which is again a contradiction. Let us now define

$$C = \{j \mid \{j\} = P_i \cap V, i \in I\} \ (\subseteq V).$$

Then this set C is a vertex cover, since for every $a^{(h,l)} \in \widetilde{T}^{\alpha}$, there exists a term t_i such that $P_i \cap V = \{h\}$ or $\{l\}$. Hence $\varphi \equiv \varphi_C$ holds for some vertex cover $C \subseteq V$, and this implies (32) by applying a discussion similar to that of (31). $\square$

5. Conclusion and discussion

In this paper, we extensively studied three types of extensions, consistent, robust and most robust, for partially defined Boolean functions that contain missing bits. In Tables 1 and 2, we summarize their complexity results for the function classes $\mathcal{C}$ considered in this

Table 1
Complexity results for typical function classes (indicated in square brackets are where the proofs can be found)

Function classes	EXTENSION	BEST-FIT	RE	CE	MRE
			$\lvert AS(a)\rvert \leqslant 1$, $\forall a \in \widetilde{T} \cup \widetilde{F}$	$\lvert AS(a)\rvert \leqslant 1$, $\forall a \in \widetilde{T} \cup \widetilde{F}$	$\lvert AS(a)\rvert \leqslant 1$, $\forall a \in \widetilde{T} \cup \widetilde{F}$
			General case	General case	General case
General	P [11]	P [11]	P [Cor.2, Th.3]	P [Th.4]	P [Th.19]
			P [Th.3]	NPC [Th.5]	NPH [Th.5]
Positive	P [11]	P [11]	P [Cor.2, Cor.3]	P [Cor.3]	P [Cor.4]
			P [Cor.3]	P [Cor.3]	NPH [Th.20]
Threshold	P [11]	NPH [2]	P [Cor.2, Th.7]	NPC [Th.8]	NPH [Th.8]
			P [Th.7]	NPC [Th.8]	NPH [Th.8]
Horn	P [11]	NPH [11]	P [Cor.2, Th.9]	NPC [Th.10]	NPH [Th.10]
			P [Th.9]	NPC [Th.10]	NPH [Th.10]
$G(S_0, G(S_1))$-decomposable	P [8]	NPH [11]	P [Cor.2]	NPC [Th.13]	NPH [Th.13]
			co-NPC [Th.11]	NPC [Th.13]	NPH [Th.11,13]
Positive $G(S_0, G(S_1))$-decomposable	P [8]	NPH [11]	P [Cor.2, 3]	P [Cor.3]	NPH [Th.21]
			P [Cor.3]	P [Cor.3]	NPH [Th.21]

P: Polynomial, NPC: NP-complete, NPH: NP-hard, co-NPC: co-NP-complete.

paper. More comprehensive results can be found in [12], in which other function classes such as regular, unate, renamable Horn, dual-minor, dual-major, self-dual, read-once and h-term k-DNF are also discussed.

Let us note at this point that a slightly modified definition of robust extension may also deserve attention on its own right. Instead of having a robust extension f, which is common to all assignments $\alpha \in \mathbb{B}^{AS}$, we define that a pBmb $(\widetilde{T}, \widetilde{F})$ is *fully consistent* with a class $\mathcal{C}$ if the pdBf $(\widetilde{T}^{\alpha}, \widetilde{F}^{\alpha})$ has an extension f_{α} belonging to $\mathcal{C}$ for every $\alpha \in \mathbb{B}^{AS}$. This gives rise to the following problem.

Problem FC($\mathcal{C}$)
Input: A pBmb $(\widetilde{T}, \widetilde{F})$, where $\widetilde{T}, \widetilde{F} \subseteq \mathbb{M}^n$.
Question: Is $(\widetilde{T}, \widetilde{F})$ fully consistent in class $\mathcal{C}$?

Let us remark that a pBmb may be fully consistent even if it has no robust extension. As an example, let us consider the pBmb $(\widetilde{T}, \widetilde{F})$ defined by

$$\widetilde{T} = \{(*, 1, 1, 1),\ (0, 0, 0, 0)\},$$
$$\widetilde{F} = \{(1, 1, 1, 0),\ (0, 1, 0, 1),\ (0, 0, 1, 1)\},$$

Table 2
Complexity results for k-DNF classes (indicated in square brackets are where the proofs can be found)

Function classes	EXTENSION	BEST-FIT	RE	CE	MRE
			$\|AS(a)\| \leqslant 1$, $\forall a \in \widetilde{T} \cup \widetilde{F}$ General case	$\|AS(a)\| \leqslant 1$, $\forall a \in \widetilde{T} \cup \widetilde{F}$ General case	$\|AS(a)\| \leqslant 1$, $\forall a \in \widetilde{T} \cup \widetilde{F}$ General case
(Positive) k-DNF	NPC [11]	NPH [11]	NPC [Cor.1] NPH [Cor.1]	NPC [Cor.1] NPC [Cor.1]	NPH [Cor.1] NPH [Cor.1]
(Positive) 1-DNF	P [11]	NPH [11]	P [Cor.2, 3, Th.14] P [Cor.3, Th.14]	P [Cor.3, Th.16] P [Cor.3, Th.16]	NPH [Th.22] NPH [Th.22]
2-DNF	P [11]	NPH [11]	P [Cor.2, Th.14] P [Th.14]	NPC [Th.17] NPC [Th.17]	NPH [Th.17, 22] NPH [Th.17, 22]
Positive 2-DNF	P [11]	NPH [11]	P [Cor.2, 3] P [Cor.3]	P [Cor.3] P [Cor.3]	NPH [Th.22] NPH [Th.22]
k-DNF with fixed $k \geqslant 3$	P [11]	NPH [11]	P [Cor.2] co-NPC [Th.15]	NPC [Th.17] NPC [Th.17]	NPH [Th.17, 22] NPH [Th.15, 17, 22]
Positive k-DNF with fixed $k \geqslant 3$	P [11]	NPH [11]	P [Cor.2, 3] P [Cor.3]	P [Cor.3] P [Cor.3]	NPH [Th.22] NPH [Th.22]

P: Polynomial, NPC: NP-complete, NPH: NP-hard, co-NPC: co-NP-complete.

and let $\mathcal{C} = \mathcal{C}_{\mathrm{TH}}$ be the class of threshold Boolean functions. There are two possible interpretations of the one missing bit, yielding

$$T^1 = \{(1,1,1,1),(0,0,0,0)\}, \qquad T^0 = \{0,1,1,1),(0,0,0,0)\}.$$

It is easy to verify that the threshold Boolean function defined by $5x_1 - 3x_2 - 3x_3 + 2x_4 \geqslant 0$ is an extension of the pdBf $(T^1, \widetilde{F})$, and that $-5x_1 + 2x_2 + 2x_3 - 3x_4 \geqslant 0$ defines a threshold extension of $(T^0, \widetilde{F})$, Hence, this pBmb is fully consistent with $\mathcal{C}_{\mathrm{TH}}$. However, $(\widetilde{T}, \widetilde{F})$ has no robust threshold extension, since the fractional vector $(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}, \frac{2}{3})$ belongs to the convex hulls of both $\widetilde{T}$ and $\widetilde{F}$.

It should also be noted that problems RE($\mathcal{C}$) and FC($\mathcal{C}$) are equivalent for many other classes, such as general, positive, regular, Horn and k-DNF. These results are reported in [15].

Finally, although this paper was written mainly from a theoretical viewpoint, these problems had arisen from real-world applications, in which the sizes of such pBmb instances are usually large. Hence it would be important to develop fast heuristic algorithms for all the problems discussed in this paper, particularly for the NP-hard cases. An attempt in this direction is found in [10].

Acknowledgement

The authors express their gratitude to the referees whose constructive comments helped significantly improve this paper.

Appendix A. Proof of Theorem 5

Theorem 5. *Problem* CE($\mathcal{C}_{\text{all}}$) *is* NP-*complete, even if* $|AS(a)| \leqslant 2$ *holds for all* $a \in \widetilde{T} \cup \widetilde{F}$.

Proof. Given an assignment $\alpha \in \mathbb{B}^{AS}$, we can check in polynomial time if $(\widetilde{T}^{\alpha}, \widetilde{F}^{\alpha})$ has an extension in $\mathcal{C}_{\text{all}}$ (see [11]). Hence problem CE($\mathcal{C}_{\text{all}}$) belongs to NP.

Let us now consider a cubic CNF

$$\Phi = \bigwedge_{k=1}^{m} C_k,$$

$$C_k = (u_k \vee v_k \vee w_k),$$

where u_k, v_k and w_k for $k = 1, 2, \ldots, m$ are literals from set

$$L = \{x_1, \bar{x}_1, \ldots, x_n, \bar{x}_n\}.$$

The 3-SAT problem, i.e., deciding the existence of a binary vector $y \in \{0, 1\}^n$ for which $\Phi(y) = 1$, is one of the well-known NP-complete problems (see [21]). We shall associate to Φ a pBmd $(\widetilde{T}, \widetilde{F})$, as follows, which has a consistent extension in $\mathcal{C}_{\text{all}}$ if and only if the 3-SAT $\Phi = 1$ has a solution.

Let us introduce subsets $A_z = \{p_{z1}, p_{z2}\}$, $z \in L$, and $B_k = \{q_{k1}, q_{k2}, q_{k3}\}$, $k = 1, 2, \ldots, m$, such that

$$A_z \cap L = B_k \cap L = A_z \cap B_k = \emptyset, \qquad A_z \cap A_{z'} = \emptyset$$

for $z \neq z'$, and $B_i \cap B_j = \emptyset$ for $i \neq j$. Let

$$V = L \cup \left(\bigcup_{z \in L} A_z\right) \cup \left(\bigcup_{k=1}^{m} B_k\right).$$

Let us denote by $(R; S)$ the vector $v \in \mathbb{M}^V$ for which $ON(v) = R$ and

$$AS(v) = \{(v, j) \mid j \in S\}.$$

(Then $OFF(v) = V \setminus (R \cup S)$, i.e., if $S = \emptyset$, then v denotes a binary vector.)

Let us construct $\widetilde{T}, \widetilde{F} \subseteq \mathbb{M}^V$ by setting

$$\widetilde{T} = \widetilde{T}_1 \cup \widetilde{T}_2, \qquad \widetilde{F} = \widetilde{F}_1 \cup \widetilde{F}_2,$$

where

$$\begin{aligned} \widetilde{T}_1 = {} & \big\{\big(L \setminus \{x_i, \bar{x}_i\}; \{x_i, \bar{x}_i\}\big) \mid x_i \in L\big\} \\ & \cup \big\{((L \setminus \{z\}) \cup \{p_{zj}\}; \emptyset) \mid z \in L, j = 1, 2\big\}, \end{aligned}$$

$$\widetilde{T}_2 = \left\{ \begin{array}{l|l} a^k = \big((L \setminus \{u_k, \bar{u}_k, v_k, \bar{v}_k, w_k, \bar{w}_k\}) \cup \{q_{k1}\}; \{q_{k2}, q_{k3}\}\big) & \\ a^{u_k 1} = \big((L \setminus \{u_k\}) \cup A_{u_k}; \{q_{k1}\}\big) & \\ a^{u_k 2} = \big((L \setminus \{u_k, \bar{u}_k, v_k\}) \cup A_{u_k} \cup \{q_{k1}\}; \{v_k\}\big) & \\ a^{u_k 3} = \big((L \setminus \{u_k, \bar{u}_k, v_k, \bar{v}_k, w_k\}) \cup A_{u_k} \cup \{q_{k1}\}; \{w_k\}\big) & \\ a^{u_k 4} = \big((L \setminus \{u_k, \bar{u}_k, v_k, \bar{v}_k, w_k, \bar{w}_k\}) \cup \{p_{u_k 2}\} \cup \{q_{k1}\}; \{p_{u_k 1}\}\big) & \\ a^{v_k 1} = \big((L \setminus \{v_k\}) \cup \{q_{k2}\}; \{q_{k1}\}\big) & k = 1, 2, \ldots, m \\ a^{v_k 2} = \big((L \setminus \{v_k, \bar{v}_k, w_k\}) \cup \{q_{k1}, q_{k2}\}; \{w_k\}\big) & \\ a^{v_k 3} = \big((L \setminus \{v_k, \bar{v}_k, w_k, \bar{w}_k, u_k\}) \cup \{q_{k1}, q_{k2}\}; \{u_k\}\big) & \\ a^{w_k 1} = \big((L \setminus \{w_k\}) \cup \{q_{k3}\}; \{q_{k1}\}\big) & \\ a^{w_k 2} = \big((L \setminus \{w_k, \bar{w}_k, u_k\}) \cup \{q_{k1}, q_{k3}\}; \{u_k\}\big) & \\ a^{w_k 3} = \big((L \setminus \{w_k, \bar{w}_k, u_k, \bar{u}_k, v_k\}) \cup \{q_{k1}, q_{k3}\}; \{v_k\}\big) & \end{array} \right\},$$

$$\widetilde{F}_1 = \{(L; \emptyset)\} \cup \{(L \setminus \{x_i, \bar{x}_i\}; \emptyset) \mid x_i \in L\} \cup \{(L \setminus \{z\}; A_z) \mid z \in L\},$$

$$\widetilde{F}_2 = \left\{ \begin{array}{l|l} b^k = \big((L \setminus \{u_k, \bar{u}_k, v_k, \bar{v}_k, w_k, \bar{w}_k\}) \cup B_k; \emptyset\big) & \\ b^{u_k 1} = \big((L \setminus \{u_k, \bar{u}_k\}) \cup A_{u_k} \cup \{q_{k1}\}; \{\bar{u}_k\}\big) & \\ b^{u_k 2} = \big((L \setminus \{u_k, \bar{u}_k, v_k, \bar{v}_k\}) \cup A_{u_k} \cup \{q_{k1}\}; \{\bar{v}_k\}\big) & \\ b^{u_k 3} = \big((L \setminus \{u_k, \bar{u}_k, v_k, \bar{v}_k, w_k, \bar{w}_k\}) \cup A_{u_k} \cup \{q_{k1}\}; \{\bar{w}_k\}\big) & \\ b^{u_k 4} = \big((L \setminus \{u_k, \bar{u}_k, v_k, \bar{v}_k, w_k, \bar{w}_k\}) \cup \{q_{k1}\}; \{p_{u_k 2}\}\big) & \\ b^{v_k 1} = \big((L \setminus \{v_k\}); \{q_{k2}\}\big) & \\ b^{v_k 2} = \big((L \setminus \{v_k, \bar{v}_k\}) \cup \{q_{k1}, q_{k2}\}; \{\bar{v}_k\}\big) & k = 1, 2, \ldots, m \\ b^{v_k 3} = \big((L \setminus \{v_k, \bar{v}_k, w_k, \bar{w}_k\}) \cup \{q_{k1}, q_{k2}\}; \{\bar{w}_k\}\big) & \\ b^{v_k 4} = \big((L \setminus \{v_k, \bar{v}_k, w_k, \bar{w}_k, u_k, \bar{u}_k\}) \cup \{q_{k1}, q_{k2}\}; \{\bar{u}_k\}\big) & \\ b^{w_k 1} = \big((L \setminus \{w_k\}); \{q_{k3}\}\big) & \\ b^{w_k 2} = \big((L \setminus \{w_k, \bar{w}_k\}) \cup \{q_{k1}, q_{k3}\}; \{\bar{w}_k\}\big) & \\ b^{w_k 3} = \big((L \setminus \{w_k, \bar{w}_k, u_k, \bar{u}_k\}) \cup \{q_{k1}, q_{k3}\}; \{\bar{u}_k\}\big) & \\ b^{w_k 4} = \big((L \setminus \{w_k, \bar{w}_k, u_k, \bar{u}_k, v_k, \bar{v}_k\}) \cup \{q_{k1}, q_{k3}\}; \{\bar{v}_k\}\big) & \end{array} \right\}.$$

It is easy to verify that $|AS(a)| \leqslant 2$ holds for all $a \in \widetilde{T} \cup \widetilde{F}$.

To prove our claim, let us first assume that there exists a consistent extension f of $(\widetilde{T}, \widetilde{F})$, and show that Φ is satisfiable. Now $(L \setminus \{x_i, \bar{x}_i\}; \{x_i, \bar{x}_i\}) \in \widetilde{T}_1$ and $(L; \emptyset)$,

$(L \setminus \{x_i, \bar{x}_i\}; \emptyset) \in \widetilde{F}_1$ imply that either $f(L \setminus \{x_i\}; \emptyset) = 1$ or $f(L \setminus \{\bar{x}_i\}; \emptyset) = 1$ (or both) holds for each of $i = 1, 2, \ldots, n$. Let us define a binary vector $y \in \mathbb{B}^n$ by

$$y_i = \begin{cases} 1 & \text{if } f(L \setminus \{x_i\}; \emptyset) = 0, \\ 0 & \text{otherwise,} \end{cases}$$

and show that this y satisfies $\Phi(y) = 1$. By the definition of y, $y_i = 1$ (respectively, $y_i = 0$) implies $f(L \setminus \{\bar{x}_i\}; \emptyset) = 1$ (respectively, $f(L \setminus \{x_i\}; \emptyset) = 1$). Assuming that there exists a clause $C_k = (u_k \vee v_k \vee w_k)$, which is 0, we derive a contradiction.

(i) If $u_k = 0$, then $f(L \setminus \{u_k\}; \emptyset) = 1$. Therefore $((L \setminus \{u_k\}); A_{u_k}) \in \widetilde{F}_1$ and $((L \setminus \{u_k\}) \cup \{p_{u_k j}\}; \emptyset) \in \widetilde{T}_1$ for $j = 1, 2$ implying

$$f\big((L \setminus \{u_k\}) \cup A_{u_k}; \emptyset\big) = 0. \tag{A.1}$$

Let us consider the sequence

$$a^{u_k 1}\big(\in \widetilde{T}_2\big), b^{u_k 1}\big(\in \widetilde{F}_2\big), \ldots, a^{u_k 4}\big(\in \widetilde{T}_2\big), b^{u_k 4}\big(\in \widetilde{F}_2\big).$$

Eq. (A.1) and $a^{u_k 1} \in \widetilde{T}_2$ imply $f((L \setminus \{u_k\}) \cup A_{u_k} \cup \{q_{k1}\}; \emptyset) = 1$, which also yields $f((L \setminus \{u_k, \bar{u}_k\}) \cup A_{u_k} \cup \{q_{k1}\}; \emptyset) = 0$ by $b^{u_k 1} \in \widetilde{F}_2$. By applying a similar argument, we have

$$f\big((L \setminus \{u_k, \bar{u}_k, v_k, \bar{v}_k, w_k, \bar{w}_k\}) \cup \{q_{k1}\}; \emptyset\big) = 0. \tag{A.2}$$

(ii) If $v_k = 0$, then $f(L \setminus \{v_k\}; \emptyset) = 1$ must hold. Let us consider the sequence

$$b^{v_k 1}\big(\in \widetilde{F}_2\big), a^{v_k 1}\big(\in \widetilde{T}_2\big), \ldots, a^{v_k 3}\big(\in \widetilde{T}_2\big), b^{v_k 4}\big(\in \widetilde{F}_2\big).$$

Then $f(L \setminus \{v_k\}; \emptyset) = 1$ and $b^{v_k 1} \in \widetilde{F}_2$ imply $f((L \setminus \{v_k\}) \cup \{q_{k2}\}; \emptyset) = 0$, from which $f((L \setminus \{v_k\}) \cup \{q_{k1}, q_{k2}\}; \emptyset) = 1$ follows by $a^{v_k 1} \in \widetilde{T}_2$. By applying a similar argument, we have

$$f\big((L \setminus \{u_k, \bar{u}_k, v_k, \bar{v}_k, w_k, \bar{w}_k\}) \cup \{q_{k1}, q_{k2}\}; \emptyset\big) = 0. \tag{A.3}$$

(iii) If $w_k = 0$, then similarly to (ii), we have

$$f\big((L \setminus \{u_k, \bar{u}_k, v_k, \bar{v}_k, w_k, \bar{w}_k\}) \cup \{q_{k1}, q_{k3}\}; \emptyset\big) = 0. \tag{A.4}$$

The three equations (A.2)–(A.4), and the fact that $b^k \in \widetilde{F}_2$ together imply that no binary assignment to the missing bits of $a^k \in \widetilde{T}_2$ can make it a true vector of f, contradicting the fact that f is a consistent extension of $(\widetilde{T}, \widetilde{F})$.

For the converse direction, let $y^* \in \mathbb{B}^n$ be a satisfying solution to Φ, and let

$$\begin{aligned} P_0 = &\big\{\big(L \setminus \{x_i\}; \emptyset\big) \mid y_i^* = 0,\ i = 1, 2, \ldots, n\big\} \\ &\cup \big\{\big(L \setminus \{\bar{x}_i\}; \emptyset\big) \mid y_i^* = 1,\ i = 1, 2, \ldots, n\big\} \\ &\cup \big\{\big((L \setminus \{z\}) \cup \{p_{zj}\}; \emptyset\big) \mid z \in L,\ j = 1, 2\big\}. \end{aligned}$$

For each clause $C_k = (u_k \vee v_k \vee w_k)$, let us define sets P_{k1}, P_{k2} and P_{k3} as follows. If $u_k = 1$ holds for the assignment y^*, then

$$P_{k1} = \left\{ \begin{array}{l} (a^k)' = \big((L \setminus \{u_k, \bar{u}_k, v_k, \bar{v}_k, w_k, \bar{w}_k\}) \cup \{q_{k1}\}; \emptyset\big) \\ (a^{u_k 1})' = \big((L \setminus \{u_k\}) \cup A_{u_k}; \emptyset\big) \\ (a^{u_k 2})' = \big((L \setminus \{u_k, \bar{u}_k\}) \cup A_{u_k} \cup \{q_{k1}\}; \emptyset\big) \\ (a^{u_k 3})' = \big((L \setminus \{u_k, \bar{u}_k, v_k, \bar{v}_k\}) \cup A_{u_k} \cup \{q_{k1}\}; \emptyset\big) \\ (a^{u_k 4})' = \big((L \setminus \{u_k, \bar{u}_k, v_k, \bar{v}_k, w_k, \bar{w}_k\}) \cup A_{u_k} \cup \{q_{k1}\}; \emptyset\big) \end{array} \middle| \; k = 1, 2, \ldots, m \right\};$$

otherwise let

$$P_{k1} = \left\{ \begin{array}{l} (a^{u_k 1})' = \big((L \setminus \{u_k\}) \cup A_{u_k} \cup \{q_{k1}\}; \emptyset\big) \\ (a^{u_k 2})' = \big((L \setminus \{u_k, \bar{u}_k, v_k\}) \cup A_{u_k} \cup \{q_{k1}\}; \emptyset\big) \\ (a^{u_k 3})' = \big((L \setminus \{u_k, \bar{u}_k, v_k, \bar{v}_k, w_k\}) \cup A_{u_k} \cup \{q_{k1}\}; \emptyset\big) \\ (a^{u_k 4})' = \big((L \setminus \{u_k, \bar{u}_k, v_k, \bar{v}_k, w_k, \bar{w}_k\}) \cup \{p_{u_k 2}\} \cup \{q_{k1}\}; \emptyset\big) \end{array} \middle| \; k = 1, 2, \ldots, m \right\}.$$

If $v_k = 1$ holds for the assignment y^*, then

$$P_{k2} = \left\{ \begin{array}{l} (a^k)'' = \big((L \setminus \{u_k, \bar{u}_k, v_k, \bar{v}_k, w_k, \bar{w}_k\}) \cup \{q_{k1}, q_{k2}\}; \emptyset\big) \\ (a^{v_k 1})' = \big((L \setminus \{v_k\}) \cup \{q_{k2}\}; \emptyset\big) \\ (a^{v_k 2})' = \big((L \setminus \{v_k, \bar{v}_k\}) \cup \{q_{k1}, q_{k2}\}; \emptyset\big) \\ (a^{v_k 3})' = \big((L \setminus \{v_k, \bar{v}_k, w_k, \bar{w}_k\}) \cup \{q_{k1}, q_{k2}\}; \emptyset\big) \end{array} \middle| \; k = 1, 2, \ldots, m \right\};$$

otherwise,

$$P_{k2} = \left\{ \begin{array}{l} (a^{v_k 1})' = \big((L \setminus \{v_k\}) \cup \{q_{k1}, q_{k2}\}; \emptyset\big) \\ (a^{v_k 2})' = \big((L \setminus \{v_k, \bar{v}_k, w_k\}) \cup \{q_{k1}, q_{k2}\}; \emptyset\big) \\ (a^{v_k 3})' = \big((L \setminus \{v_k, \bar{v}_k, w_k, \bar{w}_k, u_k\}) \cup \{q_{k1}, q_{k2}\}; \emptyset\big) \end{array} \middle| \; k = 1, 2, \ldots, m \right\}.$$

Finally, if $w_k = 1$ holds for the assignment y^*, then let

$$P_{k3} = \left\{ \begin{array}{l} (a^k)''' = \big((L \setminus \{u_k, \bar{u}_k, v_k, \bar{v}_k, w_k, \bar{w}_k\}) \cup \{q_{k1}, q_{k3}\}; \emptyset\big) \\ (a^{w_k 1})' = \big((L \setminus \{w_k\}) \cup \{q_{k3}\}; \emptyset\big) \\ (a^{w_k 2})' = \big((L \setminus \{w_k, \bar{w}_k\}) \cup \{q_{k1}, q_{k3}\}; \emptyset\big) \\ (a^{w_k 3})' = \big((L \setminus \{w_k, \bar{w}_k, u_k, \bar{u}_k\}) \cup \{q_{k1}, q_{k3}\}; \emptyset\big) \end{array} \middle| \; k = 1, 2, \ldots, m \right\};$$

otherwise set

$$P_{k3} = \left\{ \begin{array}{l} (a^{w_k 1})' = \big((L \setminus \{w_k\}) \cup \{q_{k1}, q_{k3}\}; \emptyset\big) \\ (a^{w_k 2})' = \big((L \setminus \{w_k, \bar{w}_k, u_k\}) \cup \{q_{k1}, q_{k3}\}; \emptyset\big) \\ (a^{w_k 3})' = \big((L \setminus \{w_k, \bar{w}_k, u_k, \bar{u}_k, v_k\}) \cup \{q_{k1}, q_{k3}\}; \emptyset\big) \end{array} \middle| \; k = 1, 2, \ldots, m \right\}.$$

Let us define a function f by

$$f(a)=\begin{cases}1 & \text{if } a\in P,\\ 0 & \text{otherwise,}\end{cases}$$

where $P=P_0\cup(\bigcup_{k=1}^{m}(P_{k1}\cup P_{k2}\cup P_{k3}))$. We claim that this function f is a consistent extension of $(\widetilde{T},\widetilde{F})$.

It is easy to see that for every $a\in\widetilde{T}_1$, there exists an assignment $\alpha\in\mathbb{B}^{AS(a)}$ such that $a^\alpha\in P_0$, and for every $a\in\widetilde{T}_2\setminus\{a^k\mid k=1,2,\ldots,m\}$, there exists an assignment $\alpha\in\mathbb{B}^{AS(a)}$ such that $a^\alpha=(a)'$. Finally, since y^* satisfies $C_k=1$ for each $a^k\in\widetilde{T}_2$, at least one of $(a^k)'$, $(a^k)''$ or $(a^k)'''$ belongs to P, and hence f is a consistent extension of pBmd $(\widetilde{T},\emptyset)$.

Let us show next that f is a consistent extension of $(\emptyset;\widetilde{F})$. Let

$$\begin{aligned}Q_0=\big\{(L;\emptyset)\big\}&\cup\big\{\big(L\setminus\{x_i,\bar{x}_i\};\emptyset\big)\mid x_i\in L\big\}\\ &\cup\big\{\big((L\setminus\{x_i\})\cup A_{x_i};\emptyset\big),\big(L\setminus\{\bar{x}_i\};\emptyset\big)\mid y_i^*=0,i=1,2,\ldots,n\big\}\\ &\cup\big\{(L\setminus\{x_i\};\emptyset),\big((L\setminus\{\bar{x}_i\})\cup A_{\bar{x}_i};\emptyset\big)\mid y_i^*=1,i=1,2,\ldots,n\big\}\\ &\cup\big\{b^k\in\widetilde{F}_2\mid k=1,2,\ldots,m\big\}.\end{aligned}$$

For each clause $C_k=(u_k\vee v_k\vee w_k)$, let us define sets Q_{k1}, Q_{k2} and Q_{k3} as follows. If $u_k=1$ holds for the assignment y^*, then

$$Q_{k1}=\left\{\begin{array}{l|l}(b^{u_k1})'=\big((L\setminus\{u_k\})\cup A_{u_k}\cup\{q_{k1}\};\emptyset\big) & \\ (b^{u_k2})'=\big((L\setminus\{u_k,\bar{u}_k,v_k\})\cup A_{u_k}\cup\{q_{k1}\};\emptyset\big) & \\ (b^{u_k3})'=\big((L\setminus\{u_k,\bar{u}_k,v_k,\bar{v}_k,w_k\})\cup A_{u_k}\cup\{q_{k1}\};\emptyset\big) & k=1,2,\ldots,m\\ (b^{u_k4})'=\big((L\setminus\{u_k,\bar{u}_k,v_k,\bar{v}_k,w_k,\bar{w}_k\})\cup\{p_{u_k2}\}\cup\{q_{k1}\};\emptyset\big) & \end{array}\right\};$$

otherwise let

$$Q_{k1}=\left\{\begin{array}{l|l}(b^{u_k1})'=\big((L\setminus\{u_k,\bar{u}_k\})\cup A_{u_k}\cup\{q_{k1}\};\emptyset\big) & \\ (b^{u_k2})'=\big((L\setminus\{u_k,\bar{u}_k,v_k,\bar{v}_k\})\cup A_{u_k}\cup\{q_{k1}\};\emptyset\big) & \\ (b^{u_k3})'=\big((L\setminus\{u_k,\bar{u}_k,v_k,\bar{v}_k,w_k,\bar{w}_k\})\cup A_{u_k}\cup\{q_{k1}\};\emptyset\big) & k=1,2,\ldots,m\\ (b^{u_k4})'=\big((L\setminus\{u_k,\bar{u}_k,v_k,\bar{v}_k,w_k,\bar{w}_k\})\cup\{q_{k1}\};\emptyset\big) & \end{array}\right\}.$$

If $v_k=1$ holds for the assignment y^*, then

$$Q_{k2}=\left\{\begin{array}{l|l}(b^{v_k1})'=\big((L\setminus\{v_k\});\emptyset\big) & \\ (b^{v_k2})'=\big((L\setminus\{v_k\})\cup\{q_{k1},q_{k2}\};\emptyset\big) & \\ (b^{v_k3})'=\big((L\setminus\{v_k,\bar{v}_k,w_k\})\cup\{q_{k1},q_{k2}\};\emptyset\big) & k=1,2,\ldots,m\\ (b^{v_k4})'=\big((L\setminus\{v_k,\bar{v}_k,w_k,\bar{w}_k,u_k\})\cup\{q_{k1},q_{k2}\};\emptyset\big) & \end{array}\right\};$$

otherwise,

$$Q_{k2} = \left\{ \begin{array}{l} (b^{v_k 1})' = \big((L \setminus \{v_k\}) \cup \{q_{k2}\}; \emptyset\big) \\ (b^{v_k 2})' = \big((L \setminus \{v_k, \bar{v}_k\}) \cup \{q_{k1}, q_{k2}\}; \emptyset\big) \\ (b^{v_k 3})' = \big((L \setminus \{v_k, \bar{v}_k, w_k, \bar{w}_k\}) \cup \{q_{k1}, q_{k2}\}; \emptyset\big) \\ (b^{v_k 4})' = \big((L \setminus \{v_k, \bar{v}_k, w_k, \bar{w}_k, u_k, \bar{u}_k\}) \cup \{q_{k1}, q_{k2}\}; \emptyset\big) \end{array} \middle| \, k = 1, 2, \ldots, m \right\}.$$

Finally, if $w_k = 1$ holds for the assignment y^*, then let

$$Q_{k3} = \left\{ \begin{array}{l} (b^{w_k 1})' = \big((L \setminus \{w_k\}); \emptyset\big) \\ (b^{w_k 2})' = \big((L \setminus \{w_k\}) \cup \{q_{k1}, q_{k3}\}; \emptyset\big) \\ (b^{w_k 3})' = \big((L \setminus \{w_k, \bar{w}_k, u_k\}) \cup \{q_{k1}, q_{k3}\}; \emptyset\big) \\ (b^{w_k 4})' = \big((L \setminus \{w_k, \bar{w}_k, u_k, \bar{u}_k, v_k\}) \cup \{q_{k1}, q_{k3}\}; \emptyset\big) \end{array} \middle| \, k = 1, 2, \ldots, m \right\};$$

otherwise set

$$Q_{k3} = \left\{ \begin{array}{l} (b^{w_k 1})' = \big((L \setminus \{w_k\}) \cup \{q_{k3}\}; \emptyset\big) \\ (b^{w_k 2})' = \big((L \setminus \{w_k, \bar{w}_k\}) \cup \{q_{k1}, q_{k3}\}; \emptyset\big) \\ (b^{w_k 3})' = \big((L \setminus \{w_k, \bar{w}_k, u_k, \bar{u}_k\}) \cup \{q_{k1}, q_{k3}\}; \emptyset\big) \\ (b^{w_k 4})' = \big((L \setminus \{w_k, \bar{w}_k, u_k, \bar{u}_k, v_k, \bar{v}_k\}) \cup \{q_{k1}, q_{k3}\}; \emptyset\big) \end{array} \middle| \, k = 1, 2, \ldots, m \right\}.$$

It is easy to see that for every $b \in \widetilde{F}_1 \cup \{b^k \mid k = 1, 2, \ldots, m\}$, there exists an assignment $\alpha \in \mathbb{B}^{AS(a)}$ such that $a^\alpha \in Q_0$, and for every $a \in \widetilde{F}_2 \setminus \{b^k \mid k = 1, 2, \ldots m\}$, there exists an assignment $\alpha \in \mathbb{B}^{AS(a)}$ such that $a^\alpha = (a)'$. Hence f is a consistent extension of the pBmd $(\emptyset, \widetilde{F})$.

Finally, let $Q = Q_0 \cup \big(\bigcup_{k=1}^{m} (Q_{k1} \cup Q_{k2} \cup Q_{k3})\big)$. Then $P \cap Q = \emptyset$ holds. Therefore, by combining the above two results, we can conclude that f is a consistent extension of the pBmd $(\widetilde{T}, \widetilde{F})$. $\square$

References

[1] R. Agrawal, T. Imielinski, A. Swami, Mining association rules between sets of items in large databases, in: Proc. International Conference on Management of Data (SIGMOD-93), 1993, pp. 207–216.

[2] E. Amaldi, V. Kann, The complexity of approximability of finding maximum feasible subsystems of linear relations, Theoret. Comput. Sci. 147 (1995) 181–210.

[3] D. Angluin, Queries and concept learning, Machine Learning 2 (1988) 319–342.

[4] B. Apswall, M.F. Plass, R.E.Tarjan, A linear-time algorithm for testing the truth of certain quantified Boolean formulas, Inform. Process. Lett. 8 (1979) 121–123.

[5] R.L. Ashenhurst, The decomposition of switching functions, in: Proc. International Symposium on the Theory of Switching, Part I, Harvard University Press, Cambridge, MA, 1959, pp. 75–116.

[6] S. Ben-David, E. Dichterman, Learning with restricted focus of attention, in: Proc. 6th Annual Workshop on Computational Learning Theory, 1993, pp. 287–296.

[7] A. Blumer, A. Ehrenfeucht, D. Haussler, M.K. Warmuth, Occam's razor, Inform. Process. Lett. 24 (1987) 377–380.

[8] E. Boros, V. Gurvich, P.L. Hammer, T. Ibaraki, A. Kogan, Decompositions of partially defined Boolean functions, Discrete Appl. Math. 62 (1995) 51–75.

[9] E. Boros, P.L. Hammer, J.N. Hooker, Predicting cause-effect relationships from incomplete discrete observations, SIAM J. Discrete Math. 7 (1994) 531–543.

[10] E. Boros, P.L. Hammer, T. Ibaraki, A. Kogan, E. Mayoraz, I. Muchnik, An implementation of logical analysis of data, RUTCOR Research Report RRR 22-96, Rutgers University, Piscataway, NJ, 1996.

[11] E. Boros, T. Ibaraki, K. Makino, Error-free and best-fit extensions of a partially defined Boolean function, Inform. and Comput. 140 (1998) 254–283.

[12] E. Boros, T. Ibaraki, K. Makino, Extensions of partially defined Boolean functions with missing data, RUTCOR Research Report RRR 06-96, Rutgers University, Piscataway, NJ, 1996.

[13] E. Boros, T. Ibaraki, K. Makino, Boolean analysis of incomplete examples, in: R. Karlsson, A. Lingas (Eds.), Algorithm Theory—SWAT-96, Lecture Notes in Computer Science, Vol. 1097, Springer, Berlin, 1996, pp. 440–451.

[14] E. Boros, T. Ibaraki, K. Makino, Monotone extensions of Boolean data sets, in: M. Li, A. Maruoka (Eds.), Algorithmic Learning Theory—ALT-97, Lecture Notes in Artificial Intelligence, Vol. 1316, Springer, Berlin, 1997, pp. 440–451.

[15] E. Boros, T. Ibaraki, K. Makino, Working paper.

[16] Y. Crama, P.L. Hammer, T. Ibaraki, Cause-effect relationships and partially defined Boolean functions, Ann. Oper. Res. 16 (1988) 299–326.

[17] R. Dechter, J. Pearl, Structure identification in relational data, Artificial Intelligence 58 (1992) 237–270.

[18] L. De Raedt, Logical settings for concept-learning, Artificial Intelligence 95 (1997) 187–201.

[19] D.W. Dowling, J.H. Gallier, Linear-time algorithms for testing the satisfiability of propositional Horn formulae, J. Logic Programming 3 (1984) 267–284.

[20] L.R. Ford, D.R. Fulkerson, Flows in Networks, Princeton University Press, Princeton, NJ, 1962.

[21] M.R. Garey, D.S. Johnson, Computers and Intractability, Freeman, New York, 1979.

[22] S.A. Goldman, S.S. Kwek, S.D. Scott, Learning from examples with unspecified attribute values, in: Proc. COLT-97, 1997, pp. 231–242.

[23] R. Greiner, A. Grove, A. Kogan, Exploiting the omission of irrelevant data, Artificial Intelligence 97 (1997) 345–380.

[24] D. Kavvadias, C.H. Papadimitriou, M. Sideri, On Horn envelopes and hypergraph transversals, in: K.W. Ng et al. (Eds.), Algorithms and Computation—ISAAC-93, Lecture Notes in Computer Science, Vol. 762, Springer, Berlin, 1993, pp. 399–405.

[25] R. Khardon, D. Roth, Learning to reason with a restricted view, in: Proc. COLT-95, 1995, pp. 301–310.

[26] K. Makino, K. Hatanaka, T. Ibaraki, Horn extensions of a partially defined Boolean function, RUTCOR Research Report RRR 27-95, Rutgers University, Piscataway, NJ, 1995; also: SIAM J. Comput. (to appear).

[27] K. Makino, K. Yano, T. Ibaraki, Positive and Horn decomposability of partially defined Boolean functions, Discrete Appl. Math. 74 (1997) 251–274.

[28] O.L. Mangasarian, R. Setiono and W.H. Wolberg, Pattern recognition via linear programming: Theory and applications to medical diagnosis, in: T.F. Coleman, Y. Li (Eds.), Large-Scale Numerical Optimization, SIAM, Philadelphia, PA, 1990, pp. 22–30.

[29] H. Mannila, H. Toivonen, A.I. Verkamo, Efficient algorithms for discovering association rules, in: U.M. Fayyad, R. Uthurusamy (Eds.), AAAI Workshop on Knowledge Discovery in Database, Seattle, WA, 1994, pp. 181–192.

[30] S. Muroga, Threshold Logic and Its Applications, Wiley-Interscience, New York, 1971.

[31] P.M. Murphy, D.W. Aha, UCI Repository of machine learning databases [http://www.ics.uci.edu/~mlearn/MLRepository.html], University of California, Department of Information and Computer Science, Irvine, CA, 1994.

[32] L. Pitt, L.G. Valiant, Computational limitations on learning from examples, J. ACM 35 (1988) 965–984.

[33] J.S. Provan, M.O. Ball, Efficient recognition of matroids and 2-monotonic systems, in: R. Ringeisen, F. Roberts (Eds.), Applications of Discrete Mathematics, SIAM, Philadelphia, PA, 1988, pp. 122–134.

[34] J.R. Quinlan, Induction of decision trees, Machine Learning 1 (1986) 81–106.

[35] K.G. Ramamurthy, Coherent Structures and Simple Games, Kluwer Academic Publishers, Dordrecht, Netherlands, 1990.

[36] D. Roth, Learning to reason: The non-monotonic case, in: Proc. IJCAI-95, Montreal, Quebec, 1995, pp. 1178–1184.
[37] D. Schuurmans, R. Greiner, Learning default concepts, in: Proc. 20th Canadian Conference on Artificial Intelligence (CSCSI-94), 1994.
[38] L.G. Valiant, A theory of the learnable, Comm. ACM 27 (1984) 1134–1142.
[39] L.G. Valiant, Rationality, in: Proc. Workshop on Computational Learning Theory (COLT-95), 1995, pp. 3–14.