MATHEMATICS OF OPERATIONS RESEARCH
Vol. 00, No. 0, Xxxxxx 20xx, pp. xxx–xxx
ISSN 0364-765X | EISSN 1526-5471 |xx|0000|0xxx

DOI 10.1287/moor.xxxx.xxxx

Polynomial-Time Approximation Schemes for Maximizing Gross Substitutes Utility under Budget Constraints

Akiyoshi Shioura
Graduate School of Information Sciences, Tohoku University, Sendai 980-8579, Japan
email: shioura@dais.is.tohoku.ac.jp http://www.dais.is.tohoku.ac.jp/~shioura

We consider the maximization of a gross substitutes utility function under budget constraints. This problem naturally arises in applications such as exchange economies in mathematical economics and combinatorial auctions in (algorithmic) game theory. We show that this problem admits a polynomial-time approximation scheme (PTAS). More generally, we present a PTAS for maximizing a discrete concave function called an M$^\natural$-concave function under budget constraints. Our PTAS is based on rounding an optimal solution of a continuous relaxation problem, which is shown to be solvable in polynomial time by the ellipsoid method. We also consider the maximization of the sum of two M$^\natural$-concave functions under a single budget constraint. This problem is a generalization of the budgeted max-weight matroid intersection problem to the one with certain nonlinear objective functions. We show that this problem also admits a PTAS.

Key words: discrete concave function; submodular function; budget constraints; gross substitutes utility; polynomial-time approximation scheme

MSC2000 Subject Classification: Primary: 90C27 ; Secondary: 68Q25

OR/MS subject classification: Primary: integer programming—algorithms

History: Received: Xxxx xx, xxxx; Revised: Yyyyyy yy, yyyy and Zzzzzz zz, zzzz.

1. Introduction. We consider the problem of maximizing a nonlinear utility function under a constant number of budget (or knapsack) constraints, which is formulated as

$$\text{Maximize } f(X) \quad \text{subject to } X \in 2^N,\ c_i(X) \le B_i\ (i = 1, 2, \ldots, k), \tag{1}$$

where N is a set of n items, $f : 2^N \to \mathbb{R}$ is a nonlinear utility function[1] of a consumer (or buyer) with $f(\emptyset) = 0$, k is a positive integer, and $c_i \in \mathbb{R}^N_+$, $B_i \in \mathbb{R}_+$ $(i = 1, 2, \ldots, k)$. For a vector $a \in \mathbb{R}^N$ and a set $X \subseteq N$, we denote $a(X) = \sum_{v\in X} a(v)$. The problem (1) is a natural generalization of budgeted combinatorial optimization problems ([22, 24, 43], etc.), and naturally arises in applications such as exchange economies with indivisible objects in mathematical economics ([20, 21], etc.) and combinatorial auctions in (algorithmic) game theory ([5, 10, 25], etc.).

A function $f : 2^N \to \mathbb{R}$ is said to be *submodular* if it satisfies the following condition:

$$f(X) + f(Y) \ge f(X \cup Y) + f(X \cap Y) \qquad (\forall X, Y \in 2^N).$$

The problem (1) with a submodular objective function f is extensively discussed in the literature of combinatorial optimization, and constant-factor approximation algorithms have been proposed. Wolsey [44] considered the problem (1) with a monotone submodular f and $k = 1$, and proposed the first constant-factor approximation algorithm with the ratio $1 - e^{-\beta} \simeq 0.35$, where β satisfies $e^\beta = 2 - \beta$. Later, Sviridenko [43] improved the approximation ratio to $1 - 1/e$, which is the best possible under the assumption that P $\neq$ NP [12]. For the case of a monotone submodular f and a general constant k, Kulik et al. [22] proposed a $(1 - 1/e)$-approximation algorithm by using the approach of Calinescu et al. [6] for the submodular function maximization under a matroid constraint. For a non-monotone submodular f and a general constant k, a $(0.2 - \varepsilon)$-approximation local-search algorithm was given by Lee et al. [24]. The approximation ratio is then improved in [9, 14, 23]; the best approximation ratio so far is $1/e - \varepsilon$ recently shown by Feldman et al. [14].

[1] Monotonicity of f is not assumed throughout this paper, although utility functions are often assumed to be monotone.

Submodularity for set functions is known to be equivalent to the concept of decreasing marginal utility in mathematical economics. In this paper, we focus on a more specific subclass of decreasing marginal utilities, called *gross substitutes utilities*, and show that the problem (1) admits a polynomial-time approximation scheme (PTAS) if f is a gross substitutes utility.

Gross substitutes utilities. A *gross substitutes utility* (*GS utility*, for short) function is defined as a function $f : 2^N \to \mathbb{R}$ satisfying the following condition:

$$\forall p, q \in \mathbb{R}^N \text{ with } p \leq q,\ \forall X \in \arg\max_{U \subseteq N}\{f(U) - p(U)\},$$
$$\exists Y \in \arg\max_{U \subseteq N}\{f(U) - q(U)\} \text{ such that } \{v \in X \mid p(v){=}q(v)\} \subseteq Y,$$

where p and q represent price vectors. This condition means that a consumer still wants to get items that do not change in price after the prices on other items increase. The concept of GS utility is introduced in Kelso and Crawford [21], where the existence of a Walrasian (or competitive) equilibrium is shown in a fairly general two-sided matching model. Since then, this concept plays a central role in mathematical economics and in auction theory, and is widely used in various models such as matching, housing, and labor market (see, e.g., [1, 4, 5, 10, 17, 20, 25]). While GS utility is a sufficient condition for the existence of a Walrasian equilibrium [21], it is also a necessary condition in some sense [20]. GS utility is also related to desirable properties in the auction design (see [5, 10]); for example, an optimal allocation of items in a combinatorial auction with GS utilities can be computed in polynomial time using a value oracle for utility functions (see [3] and [25, Th. 9]; see also [32, Ch. 11] and [36] for a more general result[2]).

M$^\natural$-concave functions. Various characterizations of gross substitutes utilities are given in the literature of mathematical economics [1, 17, 20]. Among them, Fujishige and Yang [17] revealed the relationship between GS utilities and discrete concave functions called *M$^\natural$-concave functions*, which is a function on matroid independent sets. It is known that a family $\mathcal{F} \subseteq 2^N$ of matroid independent sets satisfies the following property [34]:

(B$^\natural$-EXC) $\forall X, Y \in \mathcal{F}$, $\forall u \in X \setminus Y$, at least one of (i) and (ii) holds:
(i) $X - u \in \mathcal{F},\ Y + u \in \mathcal{F}$,
(ii) $\exists v \in Y \setminus X$: $X - u + v \in \mathcal{F}$, $Y + u - v \in \mathcal{F}$,

where $X - u + v$ is a short-hand notation for $(X \setminus \{u\}) \cup \{v\}$. We consider a function $f : \mathcal{F} \to \mathbb{R}$ defined on matroid independent sets $\mathcal{F}$. A function f is said to be *M$^\natural$-concave* [34] (read "M-natural-concave") if it satisfies the following:[3]

(M$^\natural$-EXC) $\forall X, Y \in \mathcal{F}$, $\forall u \in X \setminus Y$, at least one of (i) and (ii) holds:
(i) $X - u \in \mathcal{F},\ Y + u \in \mathcal{F}$, and $f(X) + f(Y) \leq f(X - u) + f(Y + u)$,
(ii) $\exists v \in Y \setminus X$: $X - u + v \in \mathcal{F}$, $Y + u - v \in \mathcal{F}$, and $f(X) + f(Y) \leq f(X - u + v) + f(Y + u - v)$.

The concept of M$^\natural$-concave function is introduced by Murota and Shioura [34] (independently of GS utilities) as a class of discrete concave functions. It is an extension of the concept of M-concave function introduced by Murota [29, 31]. In turn, M-concave functions generalize valuated matroids introduced by Dress and Wenzel [11]. The concepts of M$^\natural$-concavity/M-concavity play primary roles in the theory of discrete convex analysis [32], which provides a framework for tractable nonlinear discrete optimization problems.

It is shown by Fujishige and Yang [17] that GS utilities are essentially equivalent to M$^\natural$-concave functions; the only difference is that M$^\natural$-concave functions are defined more generally on matroid independent sets.

THEOREM 1.1 *A function $f : 2^N \to \mathbb{R}$ defined on 2^N is a gross substitutes utility if and only if f is an M$^\natural$-concave function.*

[2] In the optimal allocation problem discussed in [3, 25], we aim at optimally allocating items to *consumers*, where each item is available by only *one unit* (see Example 2.5 for details). In contrast, in the problem discussed in [36] and [32, Ch. 11] we consider *producers* of items in addition to consumers, and allow to have *multiple units* of items. It is shown in [36] and [32, Ch. 11] that there exists an optimal allocation (equilibrium, more precisely) and such an allocation can be found efficiently under the assumptions such as the gross substitute condition of consumers' utility functions.

[3] The concept of M$^\natural$-concavity is originally introduced for a function defined on (the set of integral vectors in) an integral generalized polymatroid (see [34]). In this paper a restricted class of M$^\natural$-concave functions is considered; see Section 2.

This result initiated a strong interaction between discrete convex analysis and mathematical economics; the results obtained in discrete convex analysis are used in mathematical economics ([4, 25], etc.), while mathematical economics provides interesting applications in discrete convex analysis ([36, 37], etc.).

In this paper, we consider the *k-budgeted M$^\natural$-concave maximization problem*:

$$\textbf{(}\boldsymbol{k}\textbf{BM}^\natural\textbf{M)} \quad \text{Maximize } f(X) \quad \text{subject to } X \in \mathcal{F},\ c_i(X) \le B_i\ (i = 1, 2, \ldots, k),$$

where $f : \mathcal{F} \to \mathbb{R}$ is an M$^\natural$-concave function with $f(\emptyset) = 0$ defined on matroid independent sets $\mathcal{F}$, k is a positive integer, and $c_i \in \mathbb{R}_+^N$, $B_i \in \mathbb{R}_+$ $(i = 1, 2, \ldots, k)$. Note that the problem (kBM$^\natural$M) has an additional constraint $X \in \mathcal{F}$, and if $\mathcal{F} = 2^N$, then the problem (kBM$^\natural$M) coincides with (1). We assume that the objective function f is given by a *value oracle* which, given a subset $X \in 2^N$, checks if $X \in \mathcal{F}$ or not, and returns the value $f(X)$ if $X \in \mathcal{F}$. The class of M$^\natural$-concave functions includes, as its subclass, linear functions on matroid independent sets. Hence, the problem (kBM$^\natural$M) is a nonlinear generalization of the max-weight matroid independent set problem with budget constraints, for which Grandoni and Zenklusen [18] have proposed a simple deterministic PTAS using the polyhedral structure of matroids. Note that for a more general problem called the max-weight matroid intersection problem with budget constraints, a randomized PTAS is proposed by Chekuri et al. [8].

Remark 1.1 As mentioned above, the problem (1) with a GS utility function is a special case of the problem (kBM$^\natural$M). On the other hand, the problem (kBM$^\natural$M) can be reduced to the problem (1) with an appropriately defined GS utility function; that is, these two problems are equivalent. Indeed, given an instance of (kBM$^\natural$M) with an M$^\natural$-concave function $f : \mathcal{F} \to \mathbb{R}$, the function $\tilde{f} : 2^N \to \mathbb{R}$ given by

$$\tilde{f}(X) = \max\{f(Y) \mid Y \in \mathcal{F},\ Y \subseteq X\} \qquad (X \in 2^N) \tag{2}$$

is a GS utility function, and every *minimal* optimal solution of the problem (1) with the objective function $\tilde{f}$ is an optimal solution of (kBM$^\natural$M). See Appendix A for more details. □

Our main result. In this paper, we propose a PTAS for (kBM$^\natural$M) by extending the approach of Grandoni and Zenklusen [18]. In the following, we assume that numbers such as $c_i(j)$, B_i, and $f(X)$ in the problem (kBM$^\natural$M) are all rational numbers. For a rational number r, we denote by $\langle r \rangle$ its encoding length[4]. To describe the running time of our algorithms, we use two parameters Φ and Ψ representing the input size of the problem (kBM$^\natural$M):

$$\Phi = \max_{X \in \mathcal{F}} \langle f(X) \rangle, \qquad \Psi = \max\left[\max_{1 \le i \le k, j \in N} \langle c_i(j) \rangle,\ \max_{1 \le i \le k} \langle B_i \rangle\right]. \tag{3}$$

To obtain a PTAS for (kBM$^\natural$M), we show the following property. We may assume that the following condition holds (see Proposition 3.1 for the validity of this assumption):

$$\{v\} \text{ is a feasible solution to } (k\text{BM}^\natural\text{M}) \text{ such that } f(\{v\}) > 0 \qquad (\forall v \in N). \tag{4}$$

We denote by OPT the optimal value of (kBM$^\natural$M).

Theorem 1.2
(i) *Suppose that f is an integer-valued function. Then, a feasible solution $\tilde{X} \in 2^N$ to (kBM$^\natural$M) satisfying*

$$f(\tilde{X}) \ge \text{OPT} - 2k \max_{v \in N} f(\{v\})$$

can be computed deterministically in time polynomial in n, k, Φ, and Ψ.
(ii) *For a general f and a real number ε with $0 < \varepsilon < 1$, a feasible solution $\tilde{X} \in 2^N$ to (kBM$^\natural$M) satisfying*

$$f(\tilde{X}) \ge (1 - \varepsilon)\text{OPT} - 2k \max_{v \in N} f(\{v\})$$

can be computed deterministically in time polynomial in n, k, Φ, Ψ, and $\log(1/\varepsilon)$.

[4] For an integer h, its encoding length $\langle h \rangle$ is given by $\langle h \rangle = 1 + \lceil \log_2(|h| + 1) \rceil$; for a rational number $r = p/q$ with $p, q \in \mathbb{Z}$, its encoding length $\langle r \rangle$ is given by $\langle r \rangle = \langle p \rangle + \langle q \rangle$ (see, e.g., [19, Ch. 1] for a precise definition).

Proofs of the claims (i) and (ii) are given in Sections 3.2 and 3.3, respectively. Although the bound in the statement (ii) is slightly weaker than the bound in (i), it is sufficient to obtain a PTAS for (kBM$^\natural$M).

The algorithm used in Theorem 1.2 can be converted into a PTAS by using a standard technique called *partial enumeration*, which reduces the original problem to a family of problems with "small" elements, which is done by guessing a constant number of "large" elements contained in an optimal solution (see Appendix B; see also [2, 18, 22, 39]). Hence, we obtain the following:

THEOREM 1.3 *For every fixed positive integer k and every fixed real number ε with $0 < \varepsilon < 1$, a $(1-\varepsilon)$-approximate solution of* (kBM$^\natural$M) *can be computed deterministically in time polynomial in n, Φ, and Ψ.*

To prove Theorem 1.2, we use the following algorithm, which is a natural extension of the one in [18]:

Step 1: Construct a continuous relaxation problem of (kBM$^\natural$M).
Step 2: Compute a vertex optimal solution $\hat{x} \in [0,1]^N$ to the continuous relaxation problem.
Step 3: Round down the non-integral components of the optimal solution $\hat{x}$.

In [18], linear programming (LP) relaxation is used as a continuous relaxation of the budgeted max-weight matroid independent set problem, where it is shown that a *vertex optimal solution* (i.e., an optimal solution which is a vertex of the feasible region) of the resulting LP is nearly integral. Since the LP relaxation problem can be solved in polynomial time by the ellipsoid method, rounding down a vertex optimal solution yields a near-optimal solution of the original problem.

These techniques in [18], however, cannot be applied directly since the objective function in (kBM$^\natural$M) is nonlinear while it is linear in [18]. Indeed, our continuous relaxation problem is a *nonlinear programming problem* formulated as

$$\textbf{(CR)} \qquad \text{Maximize } \overline{f}(x) \qquad \text{subject to } x \in \overline{\mathcal{F}},\ c_i^\top x \le B_i \ (i = 1, 2, \ldots, k),$$

where $\overline{\mathcal{F}}$ $(\subseteq [0,1]^N)$ is the matroid polytope of the matroid $(N, \mathcal{F})$ and $\overline{f} : \overline{\mathcal{F}} \to \mathbb{R}$ is the concave closure of the function f (see Section 2 for definitions). Since the objective function of the continuous relaxation problem (CR) is nonlinear, there may be no optimal solution which is a vertex of the feasible region.

To extend the approach in [18], we first modify the definition of vertex optimal solution appropriately. With the new definition, we show that a vertex optimal solution of (CR) is nearly integral by using the polyhedral structure of M$^\natural$-concave functions.

We then show that if f is an M$^\natural$-concave function, then the continuous relaxation problem can be solved (almost) optimally in polynomial time by using the ellipsoid method of Grötschel et al. [19]. Note that the function $\overline{f}$ in (CR) is given implicitly, and the evaluation of the function value is still a nontrivial task. It is known that the evaluation of $\overline{f}$ is NP-hard for a monotone submodular function f [6].

To solve the problem (CR), we use the following new algorithmic property concerning the concave closure of M$^\natural$-concave functions, which is proven by making full use of conjugacy results in the theory of discrete convex analysis. For $x \in \overline{\mathcal{F}}$, we call a vector $p \in \mathbb{R}^N$ a *subgradient* of $\overline{f}$ at $x \in \overline{\mathcal{F}}$ if it satisfies

$$\overline{f}(y) - \overline{f}(x) \le p^\top (y - x) \qquad (\forall y \in \overline{\mathcal{F}}).$$

We denote by $\partial \overline{f}(x)$ the set of subgradients of $\overline{f}$ at x, i.e.,

$$\begin{aligned} \partial \overline{f}(x) &= \{p \in \mathbb{R}^N \mid \overline{f}(y) - \overline{f}(x) \le p^\top (y - x) \ (\forall y \in \overline{\mathcal{F}})\} \\ &= \left\{p \in \mathbb{R}^N \mid \overline{f}(x) - p^\top x = \max\{\overline{f}(y) - p^\top y \mid y \in \overline{\mathcal{F}}\}\right\}. \end{aligned} \tag{5}$$

THEOREM 1.4 *Let $x \in \overline{\mathcal{F}}$.*
(i) *If f is an integer-valued function, then the exact value of $\overline{f}(x)$ and a subgradient of $\overline{f}$ at $x \in \overline{\mathcal{F}}$ can be computed in time polynomial in n and Φ.*
(ii) *For a general f and a real number $\delta > 0$, a value $\eta \in \mathbb{R}$ and a vector $p \in \mathbb{R}^N$ satisfying*

$$\overline{f}(x) \le \eta \le \overline{f}(x) + \delta, \qquad \overline{f}(y) - \overline{f}(x) \le p^\top (y - x) + \delta \quad (\forall y \in \overline{\mathcal{F}})$$

can be computed in time polynomial in n, Φ, and $\log(1/\delta)$.

Proof of Theorem 1.4 is given in Section 3.1, where we devise polynomial-time "combinatorial" algorithms for computing a function value and a subgradient of $\bar{f}$. Polynomiality results in Theorem 1.4 also follow from the following known facts: by LP duality and ellipsoid method, the evaluation of the concave closure $\bar{f}$ is polynomially equivalent to implementing the "demand oracle" of f, i.e., solving the problem $\max\{f(Y) - p(Y) \mid Y \in \mathcal{F}\}$ for a given $p \in \mathbb{R}^N$ (see Remark 3.1 for more details), and the demand oracle for M$^\natural$-concave functions can be implemented to run in polynomial time (see Theorem 2.1). We show in Section 3.1 that Theorem 1.4 can be proven in a simpler way without using ellipsoid method, by the reduction to the minimization of a certain discrete convex function.

Our second result. We also consider another type of budgeted optimization problem, which we call the *budgeted M$^\natural$-concave intersection problem*:

$$\textbf{(1BM}^\natural\textbf{I)} \qquad \text{Maximize } f_1(X) + f_2(X) \quad \text{ subject to } X \in \mathcal{F}_1 \cap \mathcal{F}_2,\ c(X) \le B,$$

where $f_j : \mathcal{F}_j \to \mathbb{R}$ $(j = 1, 2)$ are M$^\natural$-concave functions with $f_j(\emptyset) = 0$ defined on matroid independent sets $\mathcal{F}_j$, $c \in \mathbb{R}^N_+$, and $B \in \mathbb{R}_+$. This is a nonlinear generalization of the budgeted max-weight matroid intersection problem. Indeed, if each f_j is a linear function on matroid independent sets $\mathcal{F}_j$, then the problem (1BM$^\natural$I) is nothing but the budgeted max-weight matroid intersection problem, for which Berger et al. [2] proposed a deterministic PTAS using Lagrangian relaxation and a novel patching operation. For the budgeted max-weight matroid intersection problem with (a constant number of) multiple budget constraints, a randomized PTAS is proposed by Chekuri et al. [8].

In this paper, we show that the approach of Berger et al. [2] can be extended to (1BM$^\natural$I).

Theorem 1.5 *For every fixed real number ε with $0 < \varepsilon < 1$, a $(1-\varepsilon)$-approximate solution of* (1BM$^\natural$I) *can be computed deterministically in strongly-polynomial time (i.e., in time polynomial in n).*

The following is the key property to prove Theorem 1.5, where OPT denotes the optimal value of (1BM$^\natural$I). As in the problem (kBM$^\natural$M), we may assume, without loss of generality, that

$$\{v\} \text{ is a feasible solution to (1BM}^\natural\text{I) such that } f_1(\{v\}) + f_2(\{v\}) > 0 \qquad (\forall v \in N).$$

Theorem 1.6 *A set $\tilde{X} \in \mathcal{F}_1 \cap \mathcal{F}_2$ satisfying*

$$f_1(\tilde{X}) + f_2(\tilde{X}) \ge \text{OPT} - 2 \cdot \max_{v \in N}\{f_1(\{v\}) + f_2(\{v\})\}, \quad c(\tilde{X}) \le B + \max_{v \in N} c(v)$$

can be computed in strongly-polynomial time.

Proof of this theorem is given in Section 4. This result, combined with the partial enumeration technique (see Appendix B), implies Theorem 1.5.

To extend the approach in [2], we use techniques in Murota [30] which are developed for M$^\natural$-concave intersection problem *without* budget constraints. An important tool for our algorithm and its analysis is a *weighted* auxiliary graph defined by local information around the current solution, while an *unweighted* auxiliary graph is used in [2]. This makes it possible, in particular, to analyze how much amount the value of the objective function changes after update of a solution.

Both of our PTASes for (kBM$^\natural$M) and (1BM$^\natural$I) are based on novel approaches in Grandoni and Zenklusen [18] and in Berger et al. [2], respectively. The adaption of these approaches in the present settings, however, are not trivial as they involve nonlinear discrete concave objective functions. The main technical contribution of this paper is to show that those previous techniques for budgeted *linear* maximization problems can be extended to budgeted *nonlinear* maximization problems by using some results in the theory of discrete convex analysis.

Organization of this paper. In Section 2, we review fundamental concepts and known results in discrete convex analysis, which will be used in the following discussion. A proof of Theorem 1.2 for the problem (kBM$^\natural$M) is given in Section 3, while Theorem 1.6 is proven in Section 4.

2. Preliminaries.

In this section we review the discrete concavity concepts called M$^\natural$-concavity and L$^\natural$-concavity; these concepts play primary role in the theory of discrete convex analysis. We also present some fundamental results concerning these discrete concavity concepts, which will be used in the following discussion.

2.1 Definitions and notation. We denote by $\mathbb{Z}_+$ (resp., by $\mathbb{R}_+$) the set of nonnegative integers (resp., nonnegative real numbers). Also, we denote $\mathbf{0} = (0, 0, \ldots, 0) \in \mathbb{Z}^N$ and $\mathbf{1} = (1, 1, \ldots, 1) \in \mathbb{Z}^N$. For $x = (x(1), x(2), \ldots, x(n)) \in \mathbb{R}^N$ and $Y \in 2^N$, denote $x(Y) = \sum_{j \in Y} x(j)$. For $X \subseteq N$ the characteristic vector of X is denoted by $\chi_X \in \{0,1\}^N$, i.e.,

$$\chi_X(j) = \begin{cases} 1 & (\text{if } j \in X), \\ 0 & (\text{otherwise}). \end{cases}$$

In particular, we denote $\chi_j = \chi_{\{j\}}$ for each $j \in N$. For a nonempty set family $\mathcal{F} \subseteq 2^N$, we denote by $\overline{\mathcal{F}} \subseteq [0,1]^N$ the convex hull of vectors $\{\chi_X \mid X \in \mathcal{F}\}$.

For a function $f : \mathcal{F} \to \mathbb{R}$ defined on a nonempty set family $\mathcal{F} \subseteq 2^N$, the *concave closure* $\overline{f} : \overline{\mathcal{F}} \to \mathbb{R}$ of f is given by

$$\overline{f}(x) = \max\left\{ \sum_{Y \in \mathcal{F}} \lambda_Y f(Y) \;\middle|\; \sum_{Y \in \mathcal{F}} \lambda_Y \chi_Y = x,\ \sum_{Y \in \mathcal{F}} \lambda_Y = 1,\ \lambda_Y \geq 0\ (Y \in \mathcal{F}) \right\} \qquad (x \in \overline{\mathcal{F}}). \tag{6}$$

By LP duality, the concave closure $\overline{f}$ is also given as follows:

$$\overline{f}(x) = \min\left\{ p^\top x + \alpha \mid p \in \mathbb{R}^N,\ \alpha \in \mathbb{R},\ p(Y) + \alpha \geq f(Y)\ (Y \in \mathcal{F}) \right\} \qquad (x \in \overline{\mathcal{F}}). \tag{7}$$

Note that for every function f, the concave closure $\overline{f}$ is a polyhedral concave function satisfying $\overline{f}(\chi_X) = f(X)$ for all $X \in \mathcal{F}$. Here, $\overline{f} : \overline{\mathcal{F}} \to \mathbb{R}$ is said to be *polyhedral concave* if the set $\{(x, \alpha) \mid x \in \overline{\mathcal{F}},\ \alpha \in \mathbb{R},\ \overline{f}(x) \geq \alpha\}$ is a polyhedron.

2.2 Matroids and polymatroids. Let $\mathcal{M} = (N, \mathcal{F})$ be a matroid with the family of independent sets $\mathcal{F}$ $(\subseteq 2^N)$. Recall that a pair $(N, \mathcal{F})$ of a finite set N and a set family $\mathcal{F}$ is a *matroid* if and only if the set family $\mathcal{F}$ is given as

$$\mathcal{F} = \{X \in 2^N \mid |X \cap Y| \leq \rho(Y)\ (Y \in 2^N)\}$$

by using a nondecreasing submodular function $\rho : 2^N \to \mathbb{Z}_+$ such that $\rho(Y) \leq |Y|$ $(Y \in 2^N)$ (see, e.g., [38, 41]); such a function ρ is called the *rank function* of $\mathcal{M}$. We note that if we are given a family $\mathcal{F}$ of matroid independent sets, then the function value $\rho(X)$ can be computed easily in a greedy way in strongly-polynomial time for every $X \in 2^N$. The matroid polytope $\overline{P}(\mathcal{M})$ is defined as $\overline{P}(\mathcal{M}) = \overline{\mathcal{F}}$, i.e., the convex hull of vectors $\{\chi_X \mid X \in \mathcal{F}\}$; it is also given in terms of rank function as

$$\overline{P}(\mathcal{M}) = \{x \in \mathbb{R}^N_+ \mid x(Y) \leq \rho(Y)\ (Y \in 2^N)\}.$$

A *generalized polymatroid* (*g-polymatroid*, for short) [15] is a polyhedron

$$Q = \{x \in \mathbb{R}^N \mid \mu(X) \leq x(X) \leq \rho(X)\ (X \in 2^N)\}$$

given by a pair of submodular/supermodular functions $\rho : 2^N \to \mathbb{R} \cup \{+\infty\}$, $\mu : 2^N \to \mathbb{R} \cup \{-\infty\}$ satisfying the inequality

$$\rho(X) - \mu(Y) \geq \rho(X \setminus Y) - \mu(Y \setminus X) \qquad (\forall X, Y \in 2^N).$$

If ρ and μ are integer-valued, then Q is an integral polyhedron; in such a case, we say that Q is an *integral g-polymatroid*.

2.3 M$^\natural$-concave functions. We review the definition of M$^\natural$-concavity and show some fundamental properties and examples.

Let $\mathcal{F}$ be a family of independent sets of a matroid. A function $f : \mathcal{F} \to \mathbb{R}$ is said to be *M$^\natural$-concave* if it satisfies the condition (M$^\natural$-EXC). The concept of M$^\natural$-concavity is originally introduced for functions defined on integer lattice points (see, e.g., [32]), and the present definition of M$^\natural$-concavity for set functions can be obtained by specializing the original definition through the one-to-one correspondence between set functions and functions defined on $\{0,1\}$-vectors.

M$^\natural$-concave functions have various desirable properties as discrete concavity. Global optimality is characterized by local optimality, which implies the validity of a greedy algorithm for M$^\natural$-concave function maximization. Maximization of an M$^\natural$-concave function can be done efficiently (see, e.g., [32, 34]).

THEOREM 2.1 *For an* $\mathrm{M}^\natural$*-concave function* $f : \mathcal{F} \to \mathbb{R}$ *defined on a family* $\mathcal{F} \subseteq 2^N$ *of matroid independent sets, a maximizer of* f *can be computed in* $\mathrm{O}(n^2)$ *time.*

Maximization of the sum of two $\mathrm{M}^\natural$-concave functions is a nonlinear generalization of the max-weight matroid intersection problem, and can be solved in strongly-polynomial time as well (see Appendix D). A budget constraint with uniform cost is equivalent to a cardinality constraint. Hence, (1BM$^\natural$M) (i.e., (kBM$^\natural$M) with $k = 1$) and (1BM$^\natural$I) with uniform cost can be solved in polynomial time as well.

It is known that every $\mathrm{M}^\natural$-concave function is a submodular function in the following sense (cf. [32]):

THEOREM 2.2 ([32, Th. 6.19]) *Let* $f : \mathcal{F} \to \mathbb{R}$ *be an* $\mathrm{M}^\natural$*-concave function defined on a family* $\mathcal{F} \subseteq 2^N$ *of matroid independent sets. For* $X, Y \in \mathcal{F}$ *with* $X \cup Y \in \mathcal{F}$*, we have*

$$f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y).$$

In particular, for $X, Y \in \mathcal{F}$ *with* $X \subseteq Y$ *and* $u \in X$*, we have*

$$f(X) - f(X - u) \geq f(Y) - f(Y - u).$$

Note that the sum of an $\mathrm{M}^\natural$-concave function and a linear function is again an $\mathrm{M}^\natural$-concave function, while the sum of two $\mathrm{M}^\natural$-concave functions is not $\mathrm{M}^\natural$-concave in general.

The concept of g-polymatroid is closely related to that of $\mathrm{M}^\natural$-concavity (see [32, 35]).

THEOREM 2.3 *Let* $f : \mathcal{F} \to \mathbb{R}$ *be an* $\mathrm{M}^\natural$*-concave function defined on a family* $\mathcal{F} \subseteq 2^N$ *of matroid independent sets, and* $\overline{f} : \overline{\mathcal{F}} \to \mathbb{R}$ *be the concave closure of* f *given by* (6)*. Then, the set* $\arg\max\{\overline{f}(x) - p^\top x \mid x \in \overline{\mathcal{F}}\}$ *is an integral g-polymatroid for every* $p \in \mathbb{R}^N$*.*

We give some examples of $\mathrm{M}^\natural$-concave functions and gross substitutes (GS) utility functions. Recall that a function is GS utility if and only if it is an $\mathrm{M}^\natural$-concave function defined on 2^N (see Theorem 1.1).

A simple example of $\mathrm{M}^\natural$-concave function is a linear function $f(X) = w(X)$ $(X \in \mathcal{F})$ defined on a family $\mathcal{F} \subseteq 2^N$ of matroid independent sets, where $w \in \mathbb{R}^N$. In particular, if $\mathcal{F} = 2^N$ then f is a GS utility function. Below we give some nontrivial examples. See [32, 33] for more examples of $\mathrm{M}^\natural$-concave functions.

EXAMPLE 2.1 (WEIGHTED RANK FUNCTIONS) Let $\mathcal{I} \subseteq 2^N$ be the family of independent sets of a matroid, and $w \in \mathbb{R}^N_+$. Define a function $f : 2^N \to \mathbb{R}_+$ by

$$f(X) = \max\{w(Y) \mid Y \subseteq X,\ Y \in \mathcal{I}\} \quad (X \in 2^N),$$

which is called the *weighted rank function* [6, 7]. If $w(v) = 1$ $(v \in N)$, then f is an ordinary rank function of the matroid $(N, \mathcal{I})$. Every weighted rank function is a GS utility function [42]. □

EXAMPLE 2.2 (LAMINAR CONCAVE FUNCTIONS) Let $\mathcal{T} \subseteq 2^N$ be a laminar family, i.e., $X \cap Y = \emptyset$ or $X \subseteq Y$ or $X \supseteq Y$ holds for every $X, Y \in \mathcal{T}$. For $Y \in \mathcal{T}$, let $\varphi_Y : \mathbb{Z}_+ \to \mathbb{R}$ be a univariate concave function. Define a function $f : 2^N \to \mathbb{R}$ by

$$f(X) = \sum_{Y \in \mathcal{T}} \varphi_Y(|X \cap Y|) \quad (X \in 2^N),$$

which is called a *laminar concave function* [32, Sec. 6.3] (also called an *S-valuation* in [4]). Every laminar concave function is a GS utility function. □

EXAMPLE 2.3 (MAXIMUM-WEIGHT BIPARTITE MATCHING) Consider a bipartite graph G with two vertex sets N, J and an edge set E $(\subseteq N \times J)$, where N and J correspond to workers and jobs, respectively. An edge $(u, v) \in E$ means that worker $u \in N$ has ability to process job $v \in J$, and profit $p(u, v) \in \mathbb{R}_+$ can be obtained by assigning worker u to job v. Consider a matching between workers and jobs which maximizes the total profit, and define $\mathcal{F} \subseteq 2^N$ by

$$\mathcal{F} = \{X \subseteq N \mid \exists M : \text{matching in } G \text{ s.t. } \partial_N M = X\},$$

where $\partial_N M$ denotes the set of vertices in N covered by edges in M. It is well known that $\mathcal{F}$ is a family of independent sets in a transversal matroid. Define $f : \mathcal{F} \to \mathbb{R}$ by

$$f(X) = \max \left\{ \sum_{(u,v)\in M} p(u,v) \;\middle|\; M : \text{matching in } G \text{ s.t. } \partial_N M = X \right\} \quad (X \in \mathcal{F}).$$

Then, f is an M$^\natural$-concave function [33, Sec. 11.4.2]. In particular, if G is a complete bipartite graph, then $\mathcal{F} = 2^N$ holds, and therefore f is a GS utility function. □

Example 2.4 (M$^\natural$-concave function maximization under matroid constraint) We show that the problem of maximizing an M$^\natural$-concave function under an additional matroid constraint can be reformulated as the maximization of the sum of two M$^\natural$-concave functions.

Let $f : \mathcal{F} \to \mathbb{R}$ be an M$^\natural$-concave function defined on the family $\mathcal{F} \subseteq 2^N$ of matroid independent sets, and $\mathcal{G} \subseteq 2^N$ be another family of matroid independent sets. We consider the problem of maximizing f under the constraint given by $\mathcal{G}$:

$$\max\{f(X) \mid X \in \mathcal{F} \cap \mathcal{G}\},$$

which is equivalent to $\max\{f(X) + g(X) \mid X \in \mathcal{F} \cap \mathcal{G}\}$, where $g : \mathcal{G} \to \mathbb{R}$ is the function of $\mathcal{G}$ defined by $g(X) = 0$ $(X \in \mathcal{G})$. Since g is an M$^\natural$-concave function, the latter problem is the maximization of the sum of two M$^\natural$-concave functions. □

It should be noted that in Example 2.4, the function $f' : \mathcal{F} \cap \mathcal{G} \to \mathbb{R}$ given by $f'(X) = f(X)$ $(X \in \mathcal{F} \cap \mathcal{G})$ is *not* M$^\natural$-concave in general, even if $\mathcal{F} = 2^N$ and f is a GS utility function.

The reduction in Example 2.4 shows that the maximization of an M$^\natural$-concave function under an additional matroid constraint can be solved exactly in polynomial time. On the other hand, if the objective function is replaced with the sum of two M$^\natural$-concave functions, then the problem is NP-hard (see [32]).

Example 2.5 (Optimal allocation problem in combinatorial auction) Given a set of items N and m monotone utility functions $f_i : 2^N \to \mathbb{R}$ $(i = 1, 2, \ldots, m)$, the optimal allocation problem (also referred to as the welfare maximization problem) in combinatorial auction is formulated as follows (see, e.g., [25]):

$$\text{Maximize } \sum_{i=1}^{m} f_i(X_i) \quad \text{subject to } \{X_1, X_2, \ldots, X_m\} \text{ is a partition of } N.$$

Due to the monotonicity assumption for f_i, we may relax the condition in the constraint to the following weaker one:

$$\{X_1, X_2, \ldots, X_m\} \text{ is a subpartition of } N \text{ (i.e., } X_i \cap X_{i'} = \emptyset \text{ whenever } i \neq i').$$

We show that if each f_i is a GS utility function, then this problem can be reformulated as the maximization of the sum of two M$^\natural$-concave functions.

Suppose that each f_i in the optimal allocation problem is a GS utility function. By Example 2.4, it suffices to show that the optimal allocation problem can be reduced to the maximization of an M$^\natural$-concave function under a matroid constraint. For $i = 1, 2, \ldots, m$, let $\tilde{N}_i = \{(i,j) \mid j \in N\}$, and denote $\tilde{N} = \bigcup_{i=1}^m \tilde{N}_i$. We define a function $\tilde{f} : 2^{\tilde{N}} \to \mathbb{R}$ by

$$\tilde{f}(\tilde{X}) = \sum_{i=1}^{m} f_i(X_i),$$
$$\text{where } X_i = \{j \in N \mid (i,j) \in \tilde{X} \cap \tilde{N}_i\} \ (i = 1, 2, \ldots, m). \tag{8}$$

Then, $\tilde{f}$ is an M$^\natural$-concave function (GS utility function, in particular). For $\tilde{X} \subseteq \tilde{N}$, the set family $\{X_1, X_2, \ldots, X_m\}$ given by (8) is a subpartition of N if and only if $\tilde{X} \in \widetilde{\mathcal{G}}$, where

$$\widetilde{\mathcal{G}} = \{\tilde{Y} \subseteq \tilde{N} \mid |\tilde{Y} \cap \{(i,j) \mid 1 \leq i \leq m\}| \leq 1 \ (\forall j \in N)\}.$$

Note that $\widetilde{\mathcal{G}}$ is the family of independent sets of a partition matroid. Hence, the optimal allocation problem is reduced to the maximization of the M$^\natural$-concave function $\tilde{f}$ under the matroid constraint given by $\widetilde{\mathcal{G}}$. □

2.4 Valuated matroids. We explain the concept of valuated matroid and its equivalence with M$^\natural$-concave function. Let $\mathcal{B} \subseteq 2^N$ be the family of bases in a matroid, which is characterized by the following property (see, e.g., [32]):

(B-EXC) $\forall X, Y \in \mathcal{B}$, $\forall u \in X \setminus Y$, $\exists v \in Y \setminus X$: $X - u + v \in \mathcal{B}$, $Y + u - v \in \mathcal{B}$.

Note that $|X| = |Y|$ for every $X, Y \in \mathcal{B}$. We consider a function $g : \mathcal{B} \to \mathbb{R}$ defined on the base family $\mathcal{B}$, which is called a *valuated matroid* [11] if it satisfies the following property:

(VM) $\forall X, Y \in \mathcal{B}$, $\forall u \in X \setminus Y$, $\exists v \in Y \setminus X$:

$$X - u + v,\ Y + u - v \in \mathcal{B}, \text{ and } g(X) + g(Y) \leq g(X - u + v) + g(Y + u - v).$$

To see the equivalence between valuated matroid and M$^\natural$-concave function, we show that every M$^\natural$-concave function defined on a family of matroid independent sets can be transformed to a valuated matroid which has the same information, and vice versa. It should be noted that the equivalence shown below is just a restatement of a more general result on the equivalence between M-concavity and M$^\natural$-concavity for functions defined on integer lattice points (see [32, Sec. 6.1]), where we use the fact that valuated matroid is a special case of M-concave function.

From M$^\natural$-concave function to valuated matroid. Let $f : \mathcal{F} \to \mathbb{R}$ be an M$^\natural$-concave function defined on matroid independent sets $\mathcal{F}$. We define a valuated matroid $g : \mathcal{B} \to \mathbb{R}$ having the same information as f as follows. Let $k = \max\{|X| \mid X \in \mathcal{F}\}$. Also, let $s_1, s_2, \ldots, s_k$ be elements not in N, $S = \{s_1, s_2, \ldots, s_k\}$, and $\tilde{N} = N \cup S$. Define $\mathcal{B} \subseteq 2^{\tilde{N}}$ and a function $g : \mathcal{B} \to \mathbb{R}$ by

$$\begin{aligned} \mathcal{B} &= \{\tilde{X} \subseteq \tilde{N} \mid |\tilde{X}| = k,\ \tilde{X} \cap N \in \mathcal{F}\}, \\ g(\tilde{X}) &= f(\tilde{X} \cap N) \qquad (\tilde{X} \in \mathcal{B}). \end{aligned}$$

Then, $\mathcal{B}$ is the base family of a matroid and g is a valuated matroid; see Appendix C for a proof.

From valuated matroid to M$^\natural$-concave function. Let $g : \mathcal{B} \to \mathbb{R}$ be a valuated matroid defined on matroid bases $\mathcal{B}$. We define a function $f : \mathcal{F} \to \mathbb{R}$ as follows:

$$\mathcal{F} = \{X \subseteq N \mid \exists Y \in \mathcal{B} \text{ s.t. } X \subseteq Y\}, \qquad f(X) = \max\{g(Y) \mid Y \supseteq X,\ Y \in \mathcal{B}\} \qquad (X \in \mathcal{F}).$$

Note that the restriction of f on $\mathcal{B}$ is equal to the original function g. Since $\mathcal{B}$ is the base family of a matroid, $\mathcal{F}$ is the independent set family of a matroid (see, e.g., [38, 41]). Moreover, f is an M$^\natural$-concave function; see Appendix C for a proof.

From the transformations explained above, we see that the maximization of (the sum of) M$^\natural$-concave functions can be reduced to the maximization of (the sum of) valuated matroids, and vice versa.

2.5 L$^\natural$-convex functions. We explain the concept of L$^\natural$-convexity, which is deeply related to the concept of M$^\natural$-concavity. A function $g : \mathbb{Z}^N \to \mathbb{R} \cup \{+\infty\}$ defined on the integer lattice points is said to be *L$^\natural$-convex* if it satisfies the following inequality:

$$g(p) + g(q) \geq g((p - \lambda\mathbf{1}) \vee q) + g(p \wedge (q + \lambda\mathbf{1})) \qquad (\forall p, q \in \mathbb{Z}^N,\ \forall \lambda \in \mathbb{Z}_+),$$

where $p \vee q$ and $p \wedge q$ denote the vectors obtained by component-wise maximum and minimum of two vectors $p, q \in \mathbb{R}^n$, respectively. This inequality with $\lambda = 0$ implies the submodularity of g, in particular.

Minimization of an L$^\natural$-convex function can be solved efficiently.

Theorem 2.4 ([32]) *For an L$^\natural$-convex function $g : \mathbb{Z}^N \to \mathbb{R} \cup \{+\infty\}$ such that the set* $\mathrm{dom}_{\mathbb{Z}}\, g = \{p \in \mathbb{Z}^N \mid g(p) < +\infty\}$ *is bounded, its minimizer can be computed in time polynomial in n and* $\log \max\{\|p - q\|_\infty \mid p, q \in \mathrm{dom}_{\mathbb{Z}}\, g\}$.

L$^\natural$-convexity is also defined for polyhedral convex functions. A function $g : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ is said to be *polyhedral convex* if the set $\{(x, \alpha) \mid x \in \mathbb{R}^N,\ \alpha \in \mathbb{R},\ g(x) \leq \alpha\}$ is a polyhedron. A function $g : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ is said to be *polyhedral L$^\natural$-convex* if it is a polyhedral convex function satisfying

$$g(p) + g(q) \geq g((p - \lambda\mathbf{1}) \vee q) + g(p \wedge (q + \lambda\mathbf{1})) \qquad (\forall p, q \in \mathbb{R}^N,\ \forall \lambda \in \mathbb{R}_+).$$

The next property states the conjugacy relationship between L$^\natural$-convexity and M$^\natural$-concavity.

Theorem 2.5 ([32, 35]) *Let $f : \mathcal{F} \to \mathbb{R}$ be an M$^\natural$-concave function defined on a family $\mathcal{F} \subseteq 2^N$ of matroid independent sets. Then, the function $g : \mathbb{R}^N \to \mathbb{R}$ defined by*

$$g(p) = \max\{f(Y) - p(Y) \mid Y \in \mathcal{F}\} \qquad (p \in \mathbb{R}^N)$$

is a polyhedral L$^\natural$-convex function.

Below we present some properties of (polyhedral) L$^\natural$-convex functions which will be used in the following discussion. The next theorem shows that an L$^\natural$-convex function in integer variables can be obtained from the restriction of a polyhedral L$^\natural$-convex function.

Theorem 2.6 ([31, 32]) *Let $g : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ be a polyhedral L$^\natural$-convex function. Then, function $g_{\mathbb{Z}} : \mathbb{Z}^N \to \mathbb{R} \cup \{+\infty\}$ given by*

$$g_{\mathbb{Z}}(p) = g(p) \qquad (p \in \mathbb{Z}^N) \tag{9}$$

is an L$^\natural$-convex function if $\{p \in \mathbb{Z}^N \mid g(p) < +\infty\} \neq \emptyset$.

The next property shows that (polyhedral) L$^\natural$-convexity of a function is preserved by the restriction on an interval.

Theorem 2.7 ([31, 32]) *Let $a, b \in \mathbb{R}^N$ be vectors with $a \leq b$.*
(i) *For an L$^\natural$-convex function $g : \mathbb{Z}^N \to \mathbb{R} \cup \{+\infty\}$, the function $g_a^b : \mathbb{Z}^N \to \mathbb{R} \cup \{+\infty\}$ given by*

$$g_a^b(p) = \begin{cases} g(p) & (p \in \mathbb{Z}^N,\ a \leq p \leq b), \\ +\infty & (\textit{otherwise}) \end{cases}$$

is an L$^\natural$-convex function if $\{p \in \mathbb{Z}^N \mid a \leq p \leq b,\ g(p) < +\infty\} \neq \emptyset$.
(ii) *For a polyhedral L$^\natural$-convex function $g : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$, the function $g_a^b : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ given by*

$$g_a^b(p) = \begin{cases} g(p) & (p \in \mathbb{R}^N,\ a \leq p \leq b), \\ +\infty & (\textit{otherwise}) \end{cases}$$

is a polyhedral L$^\natural$-convex function if $\{p \in \mathbb{R}^N \mid a \leq p \leq b,\ g(p) < +\infty\} \neq \emptyset$.

The following property is so-called proximity theorem, stating that a minimizer of a polyhedral L$^\natural$-convex function and a minimizer of its restriction on $\mathbb{Z}^N$ are close to each other.

Theorem 2.8 ([28]) *Let $g : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ be a polyhedral L$^\natural$-convex function such that $\{p \in \mathbb{Z}^N \mid g(p) < +\infty\} \neq \emptyset$, and $g_{\mathbb{Z}} : \mathbb{Z}^N \to \mathbb{R} \cup \{+\infty\}$ be an L$^\natural$-convex function given by (9). For every minimizer $\hat{p}$ of $g_{\mathbb{Z}}$, there exists a minimizer p_* of g such that $\|p_* - \hat{p}\|_\infty \leq n$.*

3. PTAS for k-budgeted M$^\natural$-concave maximization. Recall that our first problem is formulated as follows:

$$\textbf{(}k\textbf{BM}^\natural\textbf{M)} \quad \text{Maximize } f(X) \quad \text{subject to } X \in \mathcal{F},\ c_i(X) \leq B_i\ (i = 1, 2, \ldots, k),$$

where $\mathcal{F} \subseteq 2^N$ is the family of independent sets of a matroid and $f : \mathcal{F} \to \mathbb{R}$ is an M$^\natural$-concave function defined on $\mathcal{F}$. We show that the problem (kBM$^\natural$M) admits a PTAS by using continuous relaxation and rounding. The continuous relaxation of (kBM$^\natural$M) used in this paper is given as follows:

$$\textbf{(CR)} \quad \text{Maximize } \overline{f}(x) \quad \text{subject to } x \in \overline{\mathcal{F}},\ c_i^\top x \leq B_i\ (i = 1, 2, \ldots, k).$$

As mentioned in Introduction, it suffices to prove Theorem 1.2, a key property to show the existence of a PTAS for (kBM$^\natural$M). We first give a proof of Theorem 1.2 (i) for the case where f is an integer-valued function in Section 3.2, and a more complicated proof for the general case (Theorem 1.2 (ii)) is given in Section 3.3. The proof for the integer-valued case is much simpler, but gives an idea of our algorithm for the general case.

Throughout this section, we assume that the condition (4) holds, i.e., for each $v \in N$, the set $\{v\}$ is a feasible solution to (kBM$^\natural$M) satisfying $f(\{v\}) > 0$. Indeed, if some element v does not satisfy this condition, then such v can be removed from N, as shown in the following property.

Proposition 3.1 *Let $v \in N$. If $\{v\}$ is not a feasible solution to* (kBM$^\natural$M) *or $f(\{v\}) \leq 0$, then there exists an optimal solution $X_* \in 2^N$ to* (kBM$^\natural$M) *with $v \notin X_*$.*

Proof. For $X, Y \in 2^N$ with $X \subseteq Y$, if Y is a feasible solution to (kBM$^\natural$M), then, X is also a feasible solution. Hence, if $\{v\}$ is not a feasible solution to (kBM$^\natural$M), then no feasible solution contains the element v; in particular, no optimal solution contains v.

We then assume that $\{v\}$ is a feasible solution to (kBM$^\natural$M) such that $f(\{v\}) \leq 0$. Let X_* be an optimal solution to (kBM$^\natural$M). If $v \notin X_*$, then we are done. Hence, we assume $v \in X_*$. Since f is an M$^\natural$-concave function, we have

$$f(\{v\}) - f(\emptyset) \geq f(X_*) - f(X_* - v)$$

by Theorem 2.2. By assumption, we have $f(\{v\}) - f(\emptyset) = f(\{v\}) \leq 0$. Hence, it holds that $f(X_* \setminus \{v\}) \geq f(X_*)$. This implies that $X_* \setminus \{v\}$ is an optimal solution to (kBM$^\natural$M) that does not contain v. □

3.1 Computing the concave closure of M$^\natural$-concave functions. In this section, we prove Theorem 1.4, stating that for the concave closure $\overline{f}$ of an M$^\natural$-concave function f, the function value and a subgradient can be computed in polynomial time. The proof is given by using conjugacy results of M$^\natural$-concave functions. The algorithms given in this section play key roles in solving the continuous relaxation problem (CR).

We define a function $g : \mathbb{R}^N \to \mathbb{R}$ by

$$g(p) = \max\{f(Y) - p(Y) \mid Y \in \mathcal{F}\} \qquad (p \in \mathbb{R}^N). \tag{10}$$

By the definition (6) of the concave closure $\overline{f}$, we have

$$g(p) = \max\{\overline{f}(y) - p^\top y \mid y \in \overline{\mathcal{F}}\} \qquad (p \in \mathbb{R}^N). \tag{11}$$

The next lemma states that the function value and a subgradient of $\overline{f}$ at a vector x can be obtained by solving a certain minimization problem.

Lemma 3.1 *For $x \in \overline{\mathcal{F}}$, we have*

$$\overline{f}(x) = \min\{p^\top x + g(p) \mid p \in \mathbb{R}^N\}, \tag{12}$$

$$\partial\overline{f}(x) = \arg\min\{p^\top x + g(p) \mid p \in \mathbb{R}^N\}. \tag{13}$$

Proof. These equations follow from known results in convex analysis and the conjugacy relationship between $\overline{f}$ and g (see, e.g., [40]). We give direct and simpler proofs below.

We first prove the formula (12) for $\overline{f}(x)$. Recall the second formula (7) for the concave closure $\overline{f}$:

$$\overline{f}(x) = \min\left\{p^\top x + \alpha \;\middle|\; p \in \mathbb{R}^N,\ \alpha \in \mathbb{R},\ p(Y) + \alpha \geq f(Y)\ (Y \in \mathcal{F})\right\} \quad (x \in \overline{\mathcal{F}}). \tag{14}$$

Since the right-hand side of (14) is a minimization problem, we may assume $\alpha = g(p)$, from which (12) follows. It is noted that the minimization problem $\min\{p^\top x + g(p) \mid p \in \mathbb{R}^N\}$ has an optimal solution since this problem is essentially equivalent to the LP in the right-hand side of (14).

To prove the formula (13), we show that $p_* \in \partial\overline{f}(x)$ holds if and only if $p_* \in \arg\min\{p^\top x + g(p) \mid p \in \mathbb{R}^N\}$. By the definition (5) of $\partial\overline{f}(x)$, we have $p_* \in \partial\overline{f}(x)$ if and only if

$$\overline{f}(x) - p_*^\top x = \max\{\overline{f}(y) - p_*^\top y \mid y \in \overline{\mathcal{F}}\} = g(p_*),$$

where the last equality is by (11). Using (12), this equation can be rewritten as

$$p_*^\top x + g(p_*) = \overline{f}(x) = \min\{p^\top x + g(p) \mid p \in \mathbb{R}^N\},$$

which is equivalent to $p_* \in \arg\min\{p^\top x + g(p) \mid p \in \mathbb{R}^N\}$. □

In the following, we show that the problem $\min\{p^\top x + g(p) \mid p \in \mathbb{R}^N\}$ in Lemma 3.1 can be solved exactly in polynomial time if f is an integer-valued function, and that an approximate solution of this problem can be computed in polynomial time for a general f; such an approximate solution gives approximate value and subgradient of $\overline{f}$, as shown later.

By definition, the evaluation of the function value $g(p)$ for a given $p \in \mathbb{R}^N$ can be done by computing the value $\max\{f(Y) - p(Y) \mid Y \in \mathcal{F}\}$, which is M$^\natural$-concave function maximization and can be solved in O(n^2) time by Theorem 2.1. It is not difficult to see that the function g is a (polyhedral) convex function in p. Moreover, M$^\natural$-concavity of f implies a nice combinatorial property of g as follows.

Lemma 3.2 *The function $g : \mathbb{R}^N \to \mathbb{R}$ given by (10) is a polyhedral L$^\natural$-convex function. Moreover, its restriction $g_{\mathbb{Z}} : \mathbb{Z}^N \to \mathbb{R}$ given by $g_{\mathbb{Z}}(x) = g(x)$ $(x \in \mathbb{Z}^N)$ is an L$^\natural$-convex function (in integer variables).*

Proof. The claims follow immediately from Theorems 2.5 and 2.6. □

The next lemma shows that there exists some subgradient of $\overline{f}$ contained in a bounded finite interval.

Lemma 3.3 *For every $x \in \overline{\mathcal{F}}$, there exists $p_* \in \partial\overline{f}(x)$ such that*

$$|p_*(v)| \le 2n \max_{X \in \mathcal{F}} |f(X)| \qquad (\forall v \in N). \tag{15}$$

Moreover, if f is an integer-valued function, then there exists such integral p_*.

Proof. By the assumption (4), the polyhedron $\overline{\mathcal{F}}$ contains the vectors $\mathbf{0}$ and χ_v for all $v \in N$, implying that the polyhedron $\overline{\mathcal{F}}$ is full-dimensional. It follows that the set

$$\{(x, \alpha) \mid x \in \overline{\mathcal{F}}, \alpha \in \mathbb{R}, \alpha \le \overline{f}(x)\}$$

is a full-dimensional polyhedron since $\bar{f}$ is a polyhedral concave function. Hence, there exists a subgradient $p_* \in \partial\overline{f}(x)$ such that the set

$$D = \{y \in \overline{\mathcal{F}} \mid \overline{f}(y) - \overline{f}(x) = p_*^\top (y - x)\}$$

is a full-dimensional polyhedron. We show that such p_* satisfies the inequality (15).

By $p_* \in \partial\overline{f}(x)$ and (5), the set D can be represented as

$$D = \arg\max\{\overline{f}(y) - p_*^\top y \mid y \in \overline{\mathcal{F}}\}.$$

Hence, D is an integral g-polymatroid by Theorem 2.3 since f is an M$^\natural$-concave function. In particular, each vertex of D is a $\{0,1\}$-vector since $D \subseteq [0,1]^N$.

Let x_0 be a $\{0,1\}$-vector which is a vertex of D. We consider the tangent cone of D at x_0, which is generated by vectors in the following set W (cf. [16, Th. 3.28]):

$$\begin{aligned} W = \{+\chi_v \mid v \in N,\ x_0 + \chi_v \in D\} &\cup \{-\chi_v \mid v \in N,\ x_0 - \chi_v \in D\} \\ &\cup \{+\chi_u - \chi_v \mid u, v \in N,\ x_0 + \chi_u - \chi_v \in D\}. \end{aligned}$$

Since D is full-dimensional, its tangent cone is also full-dimensional, which implies that W contains n linear independent vectors. Hence, the vector p_* is a (unique) solution of the system of the following linear equations, where $X_0 = \{v \in N \mid x_0(v) = 1\}$ and $q \in \mathbb{R}^N$ is a variable vector:

$$\left.\begin{array}{rl} +q(v) = \overline{f}(x_0 + \chi_v) - \overline{f}(x_0) & (= f(X_0 + v) - f(X_0)) \quad (v \in N,\ x_0 + \chi_v \in D), \\ -q(v) = \overline{f}(x_0 - \chi_v) - \overline{f}(x_0) & (= f(X_0 - v) - f(X_0)) \quad (v \in N,\ x_0 - \chi_v \in D), \\ +q(u) - q(v) = \overline{f}(x_0 + \chi_u - \chi_v) - \overline{f}(x_0) & (= f(X_0 + u - v) - f(X_0)) \\ & \qquad (u, v \in N,\ x_0 + \chi_u - \chi_v \in D). \end{array}\right\} \tag{16}$$

Recall that for every $X \in \mathcal{F}$ we have $f(X) = \overline{f}(\chi_X)$.

We show that the unique solution p_* of the system (16) of linear equations is integral if f is an integer-valued function. For this, we define a directed graph $G = (V, A)$ as follows: the node set V is given by $\{r\} \cup N$, where r is an element not in N, and the arc set A is given as

$$\begin{aligned} A = \{(r, v) \mid v \in N,\ x_0 + \chi_v \in D\} &\cup \{(v, r) \mid v \in N,\ x_0 - \chi_v \in D\} \\ &\cup \{(v, u) \mid u, v \in N,\ x_0 + \chi_u - \chi_v \in D\}. \end{aligned}$$

Then, the coefficient matrix of the system (16) is a submatrix of the incidence matrix of G obtained by removing the row corresponding to the node r. Recall that the incidence matrix of a directed graph is totally unimodular (see, e.g., [41, Th. 13.9]), and a submatrix of a totally unimodular matrix is also totally unimodular. Hence, the coefficient matrix of the system (16) is totally unimodular, and therefore the system (16) has an integral solution (i.e., $p_* \in \mathbb{Z}^n$) if f is an integer-valued function.

We finally derive the inequality (15). From (16) follows that

$$|p_*(v)| \le 2 \max_{X \in \mathcal{F}} |f(X)| \quad (v \in N,\ x_0 + \chi_v \in D \text{ or } x_0 - \chi_v \in D), \tag{17}$$

$$|p_*(u) - p_*(v)| \le 2 \max_{X \in \mathcal{F}} |f(X)| \quad (u, v \in N,\ x_0 + \chi_u - \chi_v \in D). \tag{18}$$

Since the coefficient matrix of the system (16) has rank $n = |V| - 1$, the incidence matrix of G also has rank $|V| - 1$. Hence, the directed graph G is weakly connected, i.e., the undirected graph obtained by replacing all directed arcs in G with undirected ones is connected. Therefore, for every node v in G, there exists a sequence of nodes $v_0 = v, v_1, v_2, \ldots, v_h, v_{h+1} = r$ such that $(v_j, v_{j+1}) \in A$ or $(v_{j+1}, v_j) \in A$ holds for all $j = 0, 1, \ldots, h$, where $h + 1 \leq |V| - 1 = n$. By (17) and (18), it holds that

$$\begin{aligned} |p_*(v)| &\leq |p_*(v_0) - p_*(v_1)| + |p_*(v_1) - p_*(v_2)| + \cdots + |p_*(v_{k-1}) - p_*(v_h)| + |p_*(v_h)| \\ &\leq 2(h+1) \max_{X \in \mathcal{F}} |f(X)| \ \leq \ 2n \max_{X \in \mathcal{F}} |f(X)|. \end{aligned}$$

□

Below we give a proof of Theorem 1.4. We denote $\gamma = \max_{X \in \mathcal{F}} |f(X)|$; note that $\log \gamma = \mathrm{O}(\Phi)$ holds by the definition of Φ in (3).

Suppose that f is an integer-valued function. By Lemmas 3.1 and 3.3, there exists an optimal solution p_* to the problem $\min\{p^\top x + g(p) \mid p \in \mathbb{R}^N\}$ such that

$$p_* \in \mathbb{Z}^N, \qquad |p_*(v)| \leq 2n\gamma \quad (v \in N).$$

Hence, our problem can be reduced to

$$\min\{p^\top x + g_\mathbb{Z}(p) \mid p \in \mathbb{Z}^N,\ |p(v)| \leq 2n\gamma \ (v \in N)\},$$

where the function $g_\mathbb{Z} : \mathbb{Z}^N \to \mathbb{R}$ is given by $g_\mathbb{Z}(x) = g(x)$ $(x \in \mathbb{Z}^N)$. This problem is the minimization of an L$^\natural$-convex function by Theorem 2.7 and Lemma 3.2. Therefore, its optimal solution can be computed in time polynomial in n and $\log \gamma$ by Theorem 2.4. This concludes the proof of Theorem 1.4 (i).

We then consider the case of general f, and prove Theorem 1.4 (ii); i.e., we show that a value $\eta \in \mathbb{R}$ and a vector $\hat{p} \in \mathbb{R}^N$ satisfying

$$\overline{f}(x) \leq \eta \leq \overline{f}(x) + \delta, \tag{19}$$

$$\overline{f}(y) - \overline{f}(x) \leq \hat{p}^\top (y - x) + \delta \quad (\forall y \in \overline{\mathcal{F}}) \tag{20}$$

can be computed in time polynomial in n, $\log \gamma$, and $\log(1/\delta)$. The next lemma shows that such η and $\hat{p}$ can be computed easily if we obtain a vector which is sufficiently close to an optimal solution of the problem $\min\{p^\top x + g(p) \mid p \in \mathbb{R}^N\}$.

Lemma 3.4 *Let $\hat{p} \in \mathbb{R}^N$ be a vector satisfying the condition that*

$$||\hat{p} - p_*||_\infty \leq \delta/n \quad \textit{for some} \quad p_* \in \arg\min\{p^\top x + g(p) \mid p \in \mathbb{R}^N\}. \tag{21}$$

Then, the vector $\hat{p}$ satisfies (20) *and the value $\eta = \hat{p}^\top x + g(p)$ satisfies* (19).

Proof. We have $p_* \in \partial \overline{f}(x)$ by (13) in Lemma 3.1. For $y \in \overline{\mathcal{F}}$ $(\subseteq [0,1]^N)$, it holds that

$$\begin{aligned} \overline{f}(y) - \overline{f}(x) &\leq p_*^\top (y - x) \\ &= \hat{p}^\top (y - x) + (p_* - \hat{p})^\top (y - x) \\ &\leq \hat{p}^\top (y - x) + \|p_* - \hat{p}\|_\infty \sum_{i \in N} |y(i) - x(i)| \ \leq \ \hat{p}^\top (y - x) + \delta, \end{aligned}$$

where the first inequality is by $p_* \in \partial \overline{f}(x)$ and the last inequality by $\|p_* - \hat{p}\|_\infty \leq \delta/n$ and $|y(i) - x(i)| \leq 1$ for $i \in N$. Hence, (20) holds.

We then prove (19). Since $p_* \in \arg\min\{p^\top x + g(p) \mid p \in \mathbb{R}^N\}$, we have

$$\eta = \hat{p}^\top x + g(\hat{p}) \geq p_*^\top x + g(p_*) = \overline{f}(x),$$

where the last equality follows from (12) in Lemma 3.1. Let $Y \in \mathcal{F}$ be a set with $f(Y) - \hat{p}(Y) = g(\hat{p})$. Then,

$$\begin{aligned} \eta = \hat{p}^\top x + g(\hat{p}) &= \hat{p}^\top x + f(Y) - \hat{p}(Y) \\ &= p_*^\top x + f(Y) - p_*(Y) + (\hat{p} - p_*)^\top (x - \chi_Y) \\ &\leq p_*^\top x + (f(Y) - p_*(Y)) + \|\hat{p} - p_*\|_\infty \sum_{i \in N} |x(i) - \chi_Y(i)| \ \leq \ p_*^\top x + g(p_*) + \delta, \end{aligned}$$

where the last inequality is by the definition of g, $||\hat{p} - p_*||_\infty \leq \delta/n$, and $|x(i) - \chi_Y(i)| \leq 1$ $(i \in N)$. □

The next property shows that $\hat{p}$ in Lemma 3.4 can be computed by solving the following problem:

$$\min\{p^\top x + g(p) \mid p \in (\delta/n^2)\mathbb{Z}^N,\ |p(v)| \leq 2n\gamma\ (v \in N)\}, \tag{22}$$

where $(\delta/n^2)\mathbb{Z}^N$ denotes the set of vectors with each component being an integer multiple of δ/n^2.

Lemma 3.5
(i) *Every optimal solution $\hat{p}$ to the problem* (22) *satisfies the condition* (21).
(ii) *An optimal solution to the problem* (22) *can be obtained in time polynomial in n, $\log\gamma$, and* $\log(1/\delta)$.

Proof. [Proof of (i)] Let $\hat{p}$ be an optimal solution to the problem (22). By Lemma 3.3, it suffices to show that $\hat{p}$ satisfies the condition that

$$||\hat{p} - p_*||_\infty \leq \delta/n \quad \text{for some} \quad p_* \in \arg\min\{p^\top x + g(p) \mid p \in \mathbb{R}^N,\ |p(v)| \leq 2n\gamma\ (v \in N)\}. \tag{23}$$

Define a function $h : \mathbb{R}^N \to \mathbb{R}$ defined by

$$h(q) = g((\delta/n^2)q) \qquad (q \in \mathbb{R}^N).$$

Since g is polyhedral L$^\natural$-convex, the function h is also a polyhedral L$^\natural$-convex function. We consider the following problem:

$$\min\{(\delta/n^2)q^\top x + h(q) \mid q \in \mathbb{Z}^N,\ |q(v)| \leq (2n^3/\delta)\gamma\ (v \in N)\}. \tag{24}$$

It is easy to see that this problem and the problem (22) are equivalent, and the vector $\hat{q} = (n^2/\delta)\hat{p}$ is an optimal solution to the problem (24). Hence, the condition (23) for $\hat{p}$ can be rewritten as the following condition for $\hat{q}$:

$$||\hat{q} - q_*||_\infty \leq n \ \ \text{for some} \ \ q_* \in \arg\min\{(\delta/n^2)q^\top x + h(q) \mid q \in \mathbb{R}^N,\ |q(v)| \leq (2n^3/\delta)\gamma\ (v \in N)\}. \tag{25}$$

We now show that the condition (25) holds. The restriction of h on $\mathbb{Z}^n$ is an L$^\natural$-convex function by Theorem 2.6, and therefore Theorem 2.8 implies that there exists some optimal solution $q_* \in \mathbb{R}^N$ to the continuous relaxation of (24) such that $\|q_* - \hat{q}\|_\infty \leq n$. Hence, the condition (25) holds.

[Proof of (ii)] From the discussion above, it suffices to show that the problem (24) can be solved in polynomial time. Since (24) is an L$^\natural$-convex function minimization in a bounded interval, Theorem 2.4 implies that it can be solved in time polynomial in n and $\log(2n^3/\delta)\gamma$. Hence, the claim follows. □

This concludes the proof of Theorem 1.4 (ii); recall that $\log\gamma = \mathrm{O}(\Phi)$.

Remark 3.1 Theorem 1.4 can be also proven by using the result in [19] that the strong optimization on a polyhedron is polynomially equivalent to the strong separation for the same polyhedron. Note that the proof of this equivalence in [19] is based on the ellipsoid method.

By the equation (14), the evaluation of $\overline{f}(x)$ can be done by solving an optimization problem, and it can be done in polynomial time if and only if the separation problem for the polyhedron

$$\{(p, \alpha) \in \mathbb{R}^N \times \mathbb{R} \mid p(Y) + \alpha \geq f(Y)\ (\forall Y \in \mathcal{F})\}$$

can be done in polynomial time. The separation problem can be reduced to the problem of checking the inequality $\alpha \geq \max\{f(Y) - p(Y) \mid Y \in \mathcal{F}\}$, which is solvable in polynomial time by Theorem 2.1. Hence, we obtain Theorem 1.4.

Although the approach using the ellipsoid method makes it possible to compute the exact value of $\overline{f}(x)$ and a subgradient of $\overline{f}$ at x, even in the case where f is not an integer-valued function, it has a drawback that the algorithm is not "combinatorial" and the running time is much bigger than that of the approach based on L$^\natural$-convex function minimization used in Section 3.1. □

3.2 Algorithm for integer-valued functions. We give a proof of Theorem 1.2 (i) for the case where f is an integer-valued M$^\natural$-concave function. That is, we present a deterministic algorithm for computing a feasible solution $\tilde{X}$ to (kBM$^\natural$M) satisfying

$$f(\tilde{X}) \geq \text{OPT} - 2k \max_{v \in N} f(\{v\}). \tag{26}$$

Recall that numbers such as $c_i(j)$ and B_i are assumed to be rational; this assumption is essential in the ellipsoid method [19] used in this section.

The outline of the proof is as follows. It is firstly shown that the continuous relaxation problem (CR) can be solved exactly in polynomial time; moreover, it is shown by using Theorem 1.4 (i) that a *vertex optimal solution* to (CR) can be computed in polynomial time. We call an optimal solution to (CR) a *vertex optimal solution* if it is a vertex of the set of optimal solutions to (CR); note that the set of optimal solutions to (CR) is a bounded polyhedron and therefore contains a vertex.

Lemma 3.6 *If f is an integer-valued function, then a vertex optimal solution to* (CR) *can be computed in time polynomial in n, k, Φ, and Ψ.*

Proof. Proof is given in Section 3.2.1. □

It is noted that a similar statement is shown in Shioura [42] for a *monotone* M$^\natural$-concave function defined on 2^N; we here extend the result to the case of *non-monotone* M$^\natural$-concave function defined on a subset of 2^N.

We then prove that every vertex optimal solution is nearly integral in the following sense:

Lemma 3.7 *Let $\hat{x} \in [0,1]^N$ be a vertex optimal solution to* (CR). *Then, $\hat{x}$ has at most $2k$ non-integral components.*

Proof. Proof is given in Section 3.2.2. □

Lemma 3.7 generalizes a corresponding result in [18] for the budgeted matroid independent set problem.

We finally show by using Lemma 3.7 that a feasible solution $\tilde{X}$ to (kBM$^\natural$M) satisfying (26) can be obtained by rounding down non-integral components of a vertex optimal solution to (CR).

Lemma 3.8 *Let $\hat{x} \in [0,1]^N$ be a vertex optimal solution to* (CR). *Then, the set $\tilde{X} = \{v \in N \mid \hat{x}(v) = 1\}$ is a feasible solution to* (kBM$^\natural$M) *satisfying* (26).

Proof. Proof is given in Section 3.2.3. □

From Lemmas 3.6 and 3.8 follows Theorem 1.2 (i).

3.2.1 Solving continuous relaxation. Let S_* be the set of optimal solutions to (CR); note that S_* is a bounded polyhedron. To prove Lemma 3.6, we consider the problem of finding a vertex of S_*. This problem can be solved by using the result in [19, Sec. 6.5], which implies that the ellipsoid method finds a vertex of S_* in time polynomial in n, k, Φ, and Ψ, provided that the following conditions hold:

> **(C-1)** the (strong) separation problem for the feasible region of (CR) (i.e., for a given $x \in [0,1]^N$, check if x is a feasible solution or not, and if x is not feasible, then output a hyperplane separating the feasible region and x) can be solved in polynomial time,
> **(C-2)** a subgradient of $\overline{f}$ can be computed in polynomial time.

These conditions mean that a (strong) separation oracle for S_* is available.

The condition (C-2) follows immediately from Theorem 1.4 (i). The condition (C-1) can be shown as follows. Since we can easily check the inequalities $c_i^\top x \leq B_i$, it suffices to solve the separation problem for the matroid polytope $\overline{\mathcal{F}}$, which can be done in polynomial time, provided that the rank function $\rho : 2^N \to \mathbb{Z}_+$ of the matroid $(N, \mathcal{F})$ is available (see, e.g., [19, 41]). Since we have an oracle to check in constant time whether $X \in \mathcal{F}$ or not, we can compute a function value of ρ in polynomial time (see Section 2.2). Hence, the condition (C-1) holds. This concludes the proof of Lemma 3.6.

3.2.2 Near-integrality of vertex optimal solutions. We prove Lemma 3.7. Let $\hat{x} \in [0,1]^N$ be a vertex optimal solution to (CR). Then, $\hat{x}$ is a vertex of a polyhedron given as the intersection of a set

$$Q = \arg\max\{\overline{f}(x) - p^\top x \mid x \in \overline{\mathcal{F}}\}$$

for some $p \in \mathbb{R}^N$ and the set

$$K = \{x \in \mathbb{R}^N \mid c_i^\top x \leq B_i \ (i = 1, 2, \ldots, k)\}.$$

By Theorem 2.3, the set Q is an integral g-polymatroid. Hence, the vertex $\hat{x}$ is contained in a d-dimensional face F of Q for some $d \leq k$. The statement of Lemma 3.7 follows immediately from the next property, which is a generalization of [18, Theorem 3]:

Lemma 3.9 *Let $Q \subseteq \mathbb{R}^N$ be an integral g-polymatroid and $F \subseteq Q$ be a face of Q with dimension d. Then, every $x \in F$ has at most $2d$ non-integral components.*

To prove Lemma 3.9, we use the concept of base polyhedron [16] which is deeply related to the concept of g-polymatroid. A *base polyhedron* is a polyhedron given by

$$P = \{x \in \mathbb{R}^N \mid x(X) \leq \rho(X) \ (X \subseteq N), \ x(N) = \rho(N)\}$$

with a submodular function $\rho : 2^N \to \mathbb{R} \cup \{+\infty\}$ such that $\rho(\emptyset) = 0$ and $\rho(N) < +\infty$. If ρ is integer-valued, then P is an integral polyhedron, which is called an *integral base polyhedron.* It is shown (see, e.g., [16, Sec. 3.5 (a)]) that a polyhedron $Q \subseteq \mathbb{R}^N$ is a g-polymatroid if and only if the set

$$\tilde{Q} = \{(-x(N), x) \in \mathbb{R}^{\{0\} \cup N} \mid x \in Q\}, \tag{27}$$

is a base polyhedron, where 0 is a new element not in N. Note that faces of $\tilde{Q}$ have a natural one-to-one correspondence with faces of Q, and the corresponding faces have the same dimension. Hence, Lemma 3.9 for g-polymatroids can be restated in terms of base polyhedra as follows.

Lemma 3.10 *Let $P \subseteq \mathbb{R}^N$ be an integral base polyhedron and $F \subseteq P$ be a face of dimension d. Then, every $x \in F$ has at most $2d$ non-integral components.*

Proof. Suppose that the integral base polyhedron P is associated with an integer-valued submodular function $\rho : 2^N \to \mathbb{Z} \cup \{+\infty\}$ satisfying $\rho(\emptyset) = 0$ and $\rho(N) < +\infty$. Since the dimension of F is d and every $x \in F$ satisfies $x(N) = \rho(N)$, there exist $n - d - 1$ distinct sets $Y_1, Y_2, \ldots, Y_{n-d-1} \subset N$ such that

$$F = \{x \in P \mid x(Y_j) = \rho(Y_j) \ (j = 1, 2, \ldots, n - d)\},$$

where $Y_{n-d} = N$. By a standard uncrossing argument (see, e.g., [16, 19]), we can assume that $\emptyset \neq Y_1 \subset Y_2 \subset \cdots \subset Y_{n-d} = N$ holds. Let $\hat{x} \in \mathbb{R}^N$ be an arbitrarily chosen vector in F. Putting $D_j = Y_j \setminus Y_{j-1}$ ($\neq \emptyset$) $(j = 1, 2, \ldots, n - d)$, it holds that $\hat{x}(D_j) = \rho(Y_j) - \rho(Y_{j-1}) \in \mathbb{Z}$, where $Y_0 = \emptyset$. This implies that if $|D_j| = 1$ then $\hat{x}(v) \in \mathbb{Z}$ for the unique element v in D_j. Since $|N| = n$, at least $n - 2d$ sets among $D_1, D_2, \ldots, D_{n-d}$ are singleton sets. Hence, $\hat{x}$ has at most $2d$ non-integral components. □

3.2.3 Rounding of continuous solution. We prove Lemma 3.8. Given a vertex optimal solution $\hat{x} \in [0,1]^N$ to (CR), let $\tilde{x} \in \{0,1\}^N$ be a vector obtained by rounding down the non-integral components of $\hat{x}$, i.e., $\tilde{x}(v) = 1$ if $\hat{x}(v) = 1$ and $\tilde{x}(v) = 0$ otherwise. Note that $\tilde{x}$ is the characteristic vector of $\tilde{X}$ in the statement of Lemma 3.8, and therefore satisfies $\overline{f}(\tilde{x}) = f(\tilde{X})$.

We first show that $\tilde{X}$ is a feasible solution to (kBM$^\natural$M). Since $\hat{x}$ is a vector in the matroid polytope $\overline{\mathcal{F}}$ and $\mathbf{0} \leq \tilde{x} \leq \hat{x}$, the vector $\tilde{x}$ is also in $\overline{\mathcal{F}}$. Since $\overline{\mathcal{F}} \cap \mathbb{Z}^N = \{\chi_Y \mid Y \in \mathcal{F}\}$ and $\tilde{x}$ is the characteristic vector of $\tilde{X}$, we have $\tilde{X} \in \mathcal{F}$. We also have $c_i(\tilde{X}) = c_i^\top \tilde{x} \leq c_i^\top \hat{x} \leq B_i$ for all $i = 1, \ldots, k$ since $\mathbf{0} \leq \tilde{x} \leq \hat{x}$. Hence, $\tilde{X}$ is a feasible solution to (kBM$^\natural$M).

We next show the inequality $f(\tilde{X}) \geq \textsc{opt} - 2k \max_{v \in N} f(\{v\})$. We use the following property of the concave closure $\overline{f}$ of an M$^\natural$-concave function f.

Lemma 3.11 ([32, 35, 42])
(i) *Let $x, y \in \overline{\mathcal{F}}$ be vectors with $x \leq y$, $v \in N$, and $\alpha \in \mathbb{R}_+$ be a real number such that $y + \alpha\chi_v \in \overline{\mathcal{F}}$. Then, it holds that*

$$x + \alpha\chi_v \in \overline{\mathcal{F}}, \quad \overline{f}(x + \alpha\chi_v) - \overline{f}(x) \geq \overline{f}(y + \alpha\chi_v) - \overline{f}(y).$$

(ii) *For every $v \in N$ and $\alpha \in [0,1]$, it holds that*

$$\overline{f}(\alpha\chi_v) - \overline{f}(\mathbf{0}) = \alpha\{f(\{v\}) - f(\emptyset)\}.$$

Let $u \in N$ be any element with $0 < \hat{x}(u) < 1$, and consider the vector $\hat{x} - \hat{x}(u)\chi_u$ which is obtained from $\hat{x}$ by rounding down the component $\hat{x}(u)$. It holds that

$$\begin{aligned} \overline{f}(\hat{x}) \leq \overline{f}(\hat{x} - \hat{x}(u)\chi_u) + \overline{f}(\hat{x}(u)\chi_u) - \overline{f}(\mathbf{0}) &= \overline{f}(\hat{x} - \hat{x}(u)\chi_u) + \hat{x}(u)f(\{u\}) \\ &\leq \overline{f}(\hat{x} - \hat{x}(u)\chi_u) + \max_{v \in N} f(\{v\}), \end{aligned}$$

where the first inequality is by Lemma 3.11 (i) and the equality is by Lemma 3.11 (ii). By repeated application of this argument, we obtain the inequality

$$\textsc{opt} \leq \overline{f}(\hat{x}) \leq \overline{f}(\tilde{x}) + 2k \max_{v \in N} f(\{v\}) = f(\tilde{X}) + 2k \max_{v \in N} f(\{v\});$$

recall that there exist at most $2k$ non-integral components in $\hat{x}$ by Lemma 3.7.

3.3 Algorithm for general functions. We give a proof of Theorem 1.2 (ii) for the general case where f is not necessarily an integer-valued M$^\natural$-concave function. That is, we show that for a fixed $\varepsilon > 0$, a feasible solution $\tilde{X}$ to (kBM$^\natural$M) satisfying

$$f(\tilde{X}) \geq (1 - \varepsilon)\textsc{opt} - 2k \max_{v \in N} f(\{v\}) \tag{28}$$

can be computed deterministically in time polynomial in n, k, Φ, Ψ, and $\log(1/\varepsilon)$.

We give the outline of the proof. In this case, we can compute the function value and a subgradient of $\overline{f}$ only approximately (see Theorem 1.4 (ii)). Although this makes it difficult to solve (CR) exactly in polynomial time, we can still compute an almost-optimal solution in polynomial time. We denote by $\overline{\textsc{opt}}$ the optimal value of (CR).

Lemma 3.12 *For every $\varepsilon > 0$, we can compute a feasible solution x to* (CR) *with $\overline{f}(x) \geq (1 - \varepsilon)\overline{\textsc{opt}}$ in time polynomial in n, k, Φ, Ψ, and* $\log(1/\varepsilon)$.

Proof of this lemma is given in Section 3.3.1.

Note that Lemma 3.7 concerning the near-integrality of a vertex optimal solution to (CR) still holds in the case of general f. Hence, we can compute a feasible solution $\tilde{X}$ to (kBM$^\natural$M) satisfying (26) in the same way as in Section 3.2 once a vertex optimal solution to (CR) is obtained. It is, however, difficult to compute a vertex optimal solution in this case. Instead, we will compute an almost-optimal solution which is nearly integral by using Lemma 3.12.

Lemma 3.13 *For every $\varepsilon > 0$, we can compute a feasible solution $\hat{x}$ to* (CR) *such that $\overline{f}(\hat{x}) \geq (1-\varepsilon)\overline{\textsc{opt}}$ and $\hat{x}$ has at most $2k$ non-integral components, in time polynomial in n, k, Φ, Ψ, and* $\log(1/\varepsilon)$.

A possible approach to prove Lemma 3.13 is as follows: firstly compute a feasible solution to (CR) which is sufficiently close to a vertex optimal solution, and then appropriately round up or down non-integral components of the obtained feasible solution. Although the first step in this approach can be done in the same way as in the proof of Lemma 3.6, the second step requires a careful analysis in detecting which components to round up or down.

An alternative approach we use in this paper is to find a desired feasible solution $\hat{x}$ in Lemma 3.13 in a more direct way by fixing some components of a feasible solution to (CR) to 0 or 1. This can be done by approximately solving the problem (CR) with an extra constraint $x(v) = 0$ or $x(v) = 1$. A detailed proof is given in Section 3.3.2.

We finally show that a feasible solution $\tilde{X}$ to (kBM$^\natural$M) satisfying (28) can be obtained by rounding down non-integral components of a vector $\hat{x}$ in Lemma 3.13. In the same way as in Section 3.2.3, we can show that the set $\tilde{X} = \{v \in N \mid \hat{x}(v) = 1\}$ satisfies the inequality

$$f(\tilde{X}) \geq \overline{f}(\hat{x}) - 2k \max_{v \in N} f(\{v\}).$$

Since $\overline{f}(\hat{x}) \geq (1 - \varepsilon)\overline{\textsc{opt}} \geq (1 - \varepsilon)\textsc{opt}$, the desired inequality (28) follows immediately.

3.3.1 Solving continuous relaxation approximately. We give a proof of Lemma 3.12. In the proof we use the following lemma.

LEMMA 3.14 *There exists an algorithm which, for given $\beta \in \mathbb{Q}$ and $\varepsilon' > 0$, either asserts $\beta > \overline{\text{OPT}} - \varepsilon'$ or finds a feasible solution x to* (CR) *such that $\beta \le \overline{f}(x) + \varepsilon'$, and its running time is polynomial in n, k, Φ, Ψ, $\log(1/\varepsilon')$, and the encoding length of β.*

PROOF. We prove the claim by using the ellipsoid method of Grötschel et al. [19]. Define

$$L(\beta) = \{(y, \alpha) \in [0,1]^N \times \mathbb{R} \mid y \text{ is a feasible solution to (CR)},\ \beta \le \alpha \le \overline{f}(y)\}.$$

By the result in [19, Ch. 4] on the polynomial-time equivalence between the weak optimization and the weak separation, it suffices to prove that the following weak separation problem for the set $L(\beta)$ is solvable in polynomial time:

> for given $(y, \alpha) \in [0,1]^N \times \mathbb{Q}$ and a rational number $\delta > 0$, either assert that y is a feasible solution to (CR) with $\beta \le \alpha \le \overline{f}(y) + \delta$, or find a vector $(s, \xi) \in \mathbb{Q}^N \times \mathbb{Q}$ with $\|(s, \xi)\|_\infty = 1$ such that
> $$s^\top (y' - y) + \xi(\alpha' - \alpha) \le \delta \qquad (\forall (y', \alpha') \in L(\beta)).$$

Let $(y, \alpha) \in [0,1]^N \times \mathbb{Q}$. We first check whether y is a feasible solution to (CR) or not, and if not, then output a hyperplane separating y from the feasible region of (CR). This can be done in the same way as in the case of integer-valued f (see Section 3.2.1).

Suppose that y is a feasible solution to (CR). If $\alpha < \beta$, then (y, α) is not in $L(\beta)$, and we output the vector $(s, \xi) = (\mathbf{0}, -1)$. If $\alpha \ge \beta$, then we compute an approximate value of $\overline{f}(y)$. By Theorem 1.4 (ii), we can compute in polynomial time $\eta \in \mathbb{Q}$ satisfying $\overline{f}(y) \le \eta \le \overline{f}(y) + \delta$. If $\eta \ge \alpha$, then we have $\alpha \le \overline{f}(y) + \delta$, and therefore assert that y is a feasible solution to (CR) with $\beta \le \alpha \le \overline{f}(y) + \delta$. Otherwise (i.e., $\eta < \alpha$), the vector (y, α) is not in $L(\beta)$, and we compute an "approximate" subgradient of $\overline{f}$ at y. By Theorem 1.4 (ii), we can compute in polynomial time a vector $p \in \mathbb{Q}^N$ satisfying

$$\overline{f}(y') - \overline{f}(y) \le p^\top (y' - y) + \delta \qquad (\forall y' \in \overline{\mathcal{F}}).$$

It holds that $\overline{f}(y) \le \eta < \alpha$ and $\alpha' \le \overline{f}(y')$ for all $(y', \alpha') \in L(\beta)$. Hence, we have

$$\alpha' - \alpha < \overline{f}(y') - \overline{f}(y) \le p^\top (y' - y) + \delta \qquad (\forall (y', \alpha') \in L(\beta)).$$

This shows that as the output (s, ξ) of the oracle, we can use the vector $(-p, 1)$ with each component divided by $\|(-p, 1)\|_\infty$. This concludes the proof of Lemma 3.14. □

To compute a feasible solution x to (CR) with $\overline{f}(x) \ge (1 - \varepsilon)\overline{\text{OPT}}$ in polynomial time, we use Lemma 3.14 combined with binary search with respect to β. During the binary search, we maintain an interval $[\underline{\beta}, \overline{\beta}]$ and a feasible solution $x^\bullet$ to (CR) such that

$$\underline{\beta} \le \overline{f}(x^\bullet) + \frac{\varepsilon}{3} \cdot \max_{v \in N} f(\{v\}), \qquad \overline{\beta} \ge \overline{\text{OPT}} - \frac{\varepsilon}{3} \cdot \max_{v \in N} f(\{v\}).$$

Initially, we set $\underline{\beta} = 0$, $\overline{\beta} = \sum_{v \in N} f(\{v\})$, and $x^\bullet = \mathbf{0}$; note that we have $\sum_{v \in N} f(\{v\}) \ge \overline{\text{OPT}}$ since the value $\sum_{v \in N} f(\{v\})$ is an upper bound of the function values of f and also of $\overline{f}$.

In each iteration of binary search, we use Lemma 3.14 with $\beta = (\underline{\beta} + \overline{\beta})/2$ and $\varepsilon' = (\varepsilon/3) \max_{v \in N} f(\{v\})$. If $\beta > \overline{\text{OPT}} - \varepsilon'$ holds, then we update $\overline{\beta} = \beta$, and proceed to the next iteration. If we find a feasible solution x to (CR) such that $\beta \le \overline{f}(x) + \varepsilon'$, then we update $\underline{\beta} = \beta$, $x^\bullet = x$, and proceed to the next iteration.

Suppose that $\overline{\beta} - \underline{\beta} \le \varepsilon'$ holds in some iteration. Then, it holds that

$$\overline{f}(x^\bullet) \ge \underline{\beta} - \varepsilon' \ge \overline{\beta} - 2\varepsilon' \ge \overline{\text{OPT}} - 3\varepsilon' = \overline{\text{OPT}} - \varepsilon \cdot \max_{v \in N} f(\{v\}) \ge (1 - \varepsilon)\overline{\text{OPT}};$$

note that $\max_{v \in N} f(\{v\}) \le \overline{\text{OPT}}$ since for each $v \in N$ the vector χ_v is a feasible solution to (CR) by assumption (4). Hence, the current $x^\bullet$ is a desired feasible solution to (CR). The number of iterations required by binary search is

$$\mathrm{O}\left(\log \frac{\sum_{v \in N} f(\{v\})}{(\varepsilon/3) \max_{v \in N} f(\{v\})}\right) = \mathrm{O}\left(\log \frac{3n}{\varepsilon}\right).$$

This concludes the proof of Lemma 3.12.

3.3.2 Detecting integral components. To prove Lemma 3.13, we will show that there exists a polynomial-time algorithm which finds a pair of disjoint sets $F_0, F_1 \subseteq N$ with $|F_0 \cup F_1| \geq n - 2k$ such that some $(1-\varepsilon)$-approximate solution $\hat{x}$ of (CR) satisfies $\hat{x}(v) = 0$ for $v \in F_0$ and $\hat{x}(v) = 1$ for $v \in F_1$.

For a pair of disjoint sets $S, T \in 2^N$, we denote by (CR$[S,T]$) the problem (CR) with the additional constraints that $x(v) = 0$ for $v \in S$ and $x(v) = 1$ for $v \in T$. Similarly, we denote by (P$[S,T]$) the problem (kBM$^\natural$M) with the additional constraints that $X \cap S = \emptyset$ and $T \subseteq X$. That is, (P$[S,T]$) is the problem formulated as

$$\text{Maximize} \quad f_{S,T}(X) \qquad \text{subject to} \quad X \in \mathcal{F}_{S,T},\ c_i(X) \leq B_i - c_i(T)\ (1 \leq i \leq k),$$

where $\mathcal{F}_{S,T} \subseteq 2^{N \setminus (S \cup T)}$ and $f_{S,T} : \mathcal{F}_{S,T} \to \mathbb{R}$ are given as

$$\mathcal{F}_{S,T} = \{X \subseteq N \setminus (S \cup T) \mid X \cup T \in \mathcal{F}\},$$
$$f_{S,T}(X) = f(X \cup T) - f(T) \quad (X \in \mathcal{F}_{S,T}).$$

It can be shown that $(N \setminus (S \cup T), \mathcal{F}_{S,T})$ is a matroid and $f_{S,T}$ is an M$^\natural$-convex function. Hence, (P$[S,T]$) is an instance of (kBM$^\natural$M). Note that (CR$[S,T]$) coincides with the continuous relaxation of (P$[S,T]$), which follows from the fact that $\mathcal{F}$ is the family of matroid independent sets and f is an M$^\natural$-concave function. This observation and Lemma 3.12 shown in Section 3.3.1 imply that for every $\varepsilon > 0$ we can compute $(1-\varepsilon)$-approximate solution of (CR$[S,T]$) in polynomial time. We denote by $\overline{\text{OPT}}[S,T]$ the optimal value of (CR$[S,T]$); note that $\overline{\text{OPT}} = \overline{\text{OPT}}[\emptyset, \emptyset]$.

We now explain an algorithm to compute the sets F_0 and F_1. The algorithm maintains a pair of disjoint sets $S, T \in 2^N$ and a feasible solution $\hat{x}$ to (CR$[S,T]$) satisfying the following condition:

$$\overline{f}(\hat{x}) \geq \left(1 - \frac{|S \cup T| + 1}{n+1} \cdot \varepsilon\right) \overline{\text{OPT}}. \tag{29}$$

Initially, we set $S = \emptyset$, $T = \emptyset$, and the vector $\hat{x}$ is obtained by applying Lemma 3.12 to (CR). In the following iterations, an element in $N \setminus (S \cup T)$ is repeatedly added to either S or T (and $\hat{x}$ is updated) until $|S \cup T| \geq n - 2k$ holds, as explained below.

Let $S, T, \hat{x}$ be those obtained in the previous iteration. In each iteration of the algorithm, we check whether an element $u \in N \setminus (S \cup T)$ can be added to S or T. For each $u \in N \setminus (S \cup T)$, we compute a feasible solution x_0^u to (CR$[S \cup \{u\}, T]$) and a feasible solution x_1^u to (CR$[S, T \cup \{u\}]$) such that

$$\overline{f}(x_0^u) \geq \left(1 - \frac{\varepsilon}{n+1}\right) \overline{\text{OPT}}[S \cup \{u\}, T], \qquad \overline{f}(x_1^u) \geq \left(1 - \frac{\varepsilon}{n+1}\right) \overline{\text{OPT}}[S, T \cup \{u\}].$$

Suppose that $\overline{f}(x_0^u) \geq (1 - \varepsilon/(n+1))\overline{f}(\hat{x})$ holds for some $u \in N \setminus (S \cup T)$. Then, we have

$$\begin{aligned}\overline{f}(x_0^u) &\geq \left(1 - \frac{\varepsilon}{n+1}\right)\overline{f}(\hat{x}) \\ &\geq \left(1 - \frac{\varepsilon}{n+1}\right)\left(1 - \frac{|S \cup T| + 1}{n+1} \cdot \varepsilon\right) \overline{\text{OPT}} \geq \left(1 - \frac{|S \cup T| + 2}{n+1} \cdot \varepsilon\right) \overline{\text{OPT}}.\end{aligned}$$

Hence, we add the element u to S, replace $\hat{x}$ with x_0^u, and proceed to the next iteration. Similarly, if $\overline{f}(x_1^u) \geq (1 - \varepsilon/(n+1))\overline{f}(\hat{x})$ holds for some $u \in N \setminus (S \cup T)$, then we add u to T, replace $\hat{x}$ with x_1^u, and proceed to the next iteration.

Suppose that $\max\{\overline{f}(x_0^u), \overline{f}(x_1^u)\} < (1 - \varepsilon/(n+1))\overline{f}(\hat{x})$ hold for all $u \in N \setminus (S \cup T)$. Then, we have

$$\left(1 - \frac{\varepsilon}{n+1}\right) \overline{\text{OPT}}[S \cup \{u\}, T] \leq \overline{f}(x_0^u) < \left(1 - \frac{\varepsilon}{n+1}\right)\overline{f}(\hat{x}) \leq \left(1 - \frac{\varepsilon}{n+1}\right)\overline{\text{OPT}},$$
$$\left(1 - \frac{\varepsilon}{n+1}\right) \overline{\text{OPT}}[S, T \cup \{u\}] \leq \overline{f}(x_1^u) < \left(1 - \frac{\varepsilon}{n+1}\right)\overline{f}(\hat{x}) \leq \left(1 - \frac{\varepsilon}{n+1}\right)\overline{\text{OPT}},$$

implying that

$$\max\{\overline{\text{OPT}}[S \cup \{u\}, T], \overline{\text{OPT}}[S, T \cup \{u\}]\} < \overline{\text{OPT}} \qquad (\forall u \in N \setminus (S \cup T)).$$

This means that any optimal solution of the problem (CR$[S,T]$) has no more integral component. On the other hand, the problem (CR$[S,T]$) has $n' = n - |S \cup T|$ free variables, and Lemma 3.7 applied to (CR$[S,T]$) implies that there exists an optimal solution of (CR$[S,T]$) which has at least $(n' - 2k)$ integral components. Hence, we must have $n' \leq 2k$, i.e., $|S \cup T| \geq n - 2k$ holds. By (29), the current vector $\hat{x}$ satisfies $\overline{f}(\hat{x}) \geq (1 - \varepsilon)\, \overline{\text{OPT}}$. This concludes the proof of Lemma 3.13.

4. PTAS for 1-budgeted M$^\natural$-concave intersection. We give a proof of Theorem 1.6 for (1BM$^\natural$I), i.e., we show that a set $\tilde{X} \in \mathcal{F}_1 \cap \mathcal{F}_2$ satisfying the condition

$$f_1(\tilde{X}) + f_2(\tilde{X}) \geq \text{OPT} - 2 \cdot \max_{v \in N}\{f_1(\{v\}) + f_2(\{v\})\}, \quad c(\tilde{X}) \leq B + \max_{v \in N} c(v) \tag{30}$$

can be computed in strongly-polynomial time. Recall the assumption that for all $v \in N$, the set $X = \{v\}$ is a feasible solution to (1BM$^\natural$I) and satisfies $f_1(\{v\}) + f_2(\{v\}) > 0$.

4.1 Lagrangian relaxation approach. To obtain a set $\tilde{X} \in \mathcal{F}_1 \cap \mathcal{F}_2$ satisfying the condition (30), we apply the Lagrangian relaxation approach to (1BM$^\natural$I) in a similar way as in [2, 39]. With a parameter $\lambda \in \mathbb{R}_+$ called *Lagrangian multiplier*, the Lagrangian relaxation problem of (1BM$^\natural$I) is given by

$$\textbf{(LR}(\boldsymbol{\lambda})\textbf{)} \quad \text{Maximize} \quad f_1(X) + f_2(X) + \lambda(B - c(X)) \quad \text{subject to} \quad X \in \mathcal{F}_1 \cap \mathcal{F}_2.$$

The problem (LR(λ)) is an instance of the M$^\natural$-concave intersection problem *without* budget constraint. Indeed, the function $\hat{f}_1 : \mathcal{F}_1 \to \mathbb{R}$ defined by

$$\hat{f}_1(X) = f_1(X) + \lambda(B - c(X)) \quad (X \in \mathcal{F}_1) \tag{31}$$

is an M$^\natural$-concave function, and therefore the objective function of (LR(λ)) can be regarded as the sum of two M$^\natural$-concave functions $\hat{f}_1$ and f_2.

Since an M$^\natural$-concave function can be transformed to a valuated matroid which has the same information (see Section 2.4 and Appendix C), the M$^\natural$-concave intersection problem can be reduced to the valuated matroid intersection problem discussed in [29]. Hence, the theorems and algorithms in [29] for the valuated matroid intersection problem can be applied to (LR(λ)) with slight modification. In particular, (LR(λ)) can be solved in strongly-polynomial time (see Appendix D).

We denote by $z_{\text{LR}}(\lambda)$ the optimal value of (LR(λ)), i.e.,

$$z_{\text{LR}}(\lambda) = \max\{f_1(X) + f_2(X) + \lambda(B - c(X)) \mid X \in \mathcal{F}_1 \cap \mathcal{F}_2\} \qquad (\lambda \in \mathbb{R}). \tag{32}$$

By definition, z_{LR} is a piecewise-linear convex function given as the upper envelope of many linear functions $f_1(X) + f_2(X) + \lambda(B - c(X))$ $(X \in \mathcal{F}_1 \cap \mathcal{F}_2)$, where λ is a variable. Therefore, for each interval $[\eta, \zeta]$ such that the function z_{LR} is linear in $[\eta, \zeta]$, there exists some $\hat{X} \in \mathcal{F}_1 \cap \mathcal{F}_2$ such that

$$z_{\text{LR}}(\lambda) = f_1(\hat{X}) + f_2(\hat{X}) + \lambda(B - c(\hat{X})) \qquad (\forall \lambda \in [\eta, \zeta])$$

and $\hat{X}$ is an optimal solution to the problem (LR(λ)) for every $\lambda \in [\eta, \zeta]$.

A value $\lambda = \lambda_*$ minimizing $z_{\text{LR}}(\lambda)$ is called an *optimal Lagrangian multiplier*. Since z_{LR} is a convex function in λ, an optimal Lagrangian multiplier λ_* is characterized by the condition that

$$(z_{\text{LR}})'_+(\lambda_*) \geq 0, \qquad (z_{\text{LR}})'_-(\lambda_*) \leq 0. \tag{33}$$

Here, $(z_{\text{LR}})'_+(\lambda)$ and $(z_{\text{LR}})'_-(\lambda)$ denote the left derivative and the right derivative of the convex function z_{LR} at $\lambda \in \mathbb{R}_+$, respectively, which are defined by

$$(z_{\text{LR}})'_+(\lambda) = \lim_{\lambda' \downarrow \lambda} \frac{z_{\text{LR}}(\lambda') - z_{\text{LR}}(\lambda)}{\lambda' - \lambda}, \qquad (z_{\text{LR}})'_-(\lambda) = \lim_{\lambda' \uparrow \lambda} \frac{z_{\text{LR}}(\lambda') - z_{\text{LR}}(\lambda)}{\lambda' - \lambda}.$$

The next lemma shows that left and right derivatives of z_{LR} can be computed in strongly-polynomial time.

LEMMA 4.1 *Let $\lambda \in \mathbb{R}_+$, and δ be a sufficiently small positive real number. Also, let X_* and Y_* be optimal solution to the problems* (LR($\lambda + \delta$)) *and* (LR($\lambda - \delta$))*, respectively. Then, X_* and Y_* have the minimum value of $c(X_*)$ and the maximum value of $c(Y_*)$ among all optimal solutions to* (LR(λ))*, and satisfy*

$$(z_{\text{LR}})'_+(\lambda) = B - c(X_*), \qquad (z_{\text{LR}})'_-(\lambda) = B - c(Y_*). \tag{34}$$

PROOF. We give a proof of the statement for X_* only since the statement for Y_* can be shown similarly. Since z_{LR} is a piecewise-linear function and δ is a sufficiently small number, the function z_{LR} is linear in the interval $[\lambda, \lambda + 2\delta]$. Hence, there exists some $\hat{X} \in \mathcal{F}_1 \cap \mathcal{F}_2$ such that

$$z_{\text{LR}}(\lambda') = f_1(\hat{X}) + f_2(\hat{X}) + \lambda'(B - c(\hat{X})) \qquad (\forall \lambda' \in [\lambda, \lambda + 2\delta]). \tag{35}$$

Since X_* is an optimal solution to the problem (LR($\lambda+\delta$)) and the function $z_{\rm LR}$ is linear in the interval $[\lambda, \lambda+2\delta]$, the set X_* is also an optimal solution to the problem (LR(λ')) for every $\lambda' \in [\lambda, \lambda+2\delta]$, which implies that $\hat{X} = X_*$ satisfies the equation (35). Hence, we have $(z_{\rm LR})'_+(\lambda) = B - c(X_*)$.

Suppose, to the contrary, that there exists some optimal solution X' to (LR(λ)) such that $c(X') < c(X_*)$. Since both of X' and X_* are optimal to (LR(λ)), we have

$$f_1(X') + f_2(X') + \lambda(B - c(X')) = f_1(X_*) + f_2(X_*) + \lambda(B - c(X_*)),$$

which, combined with the inequality $c(X') < c(X_*)$, implies that

$$f_1(X') + f_2(X') + (\lambda+\delta)(B - c(X')) > f_1(X_*) + f_2(X_*) + (\lambda+\delta)(B - c(X_*)),$$

a contradiction to the fact that X_* is an optimal solution to (LR($\lambda+\delta$)). Therefore, X_* minimizes the value $c(X_*)$ among all optimal solutions to (LR(λ)). □

An optimal Lagrangian multiplier can be computed in polynomial time. Indeed, since the optimality condition (33) can be checked in polynomial time by Lemma 4.1, an optimal Lagrangian multiplier can be found in weakly-polynomial time by binary search, provided that the input numbers such as $c(j)$, B, $f_1(X)$, and $f_2(X)$ are all rational numbers. Moreover, this can be done in strongly-polynomial time by using the parametric approach of Megiddo [27] in the same way as in [2, 39] (see Appendix E for details).

Lemma 4.2 *An optimal Lagrangian multiplier can be computed in time polynomial in n.*

We show some properties of optimal solutions to (LR(λ_*)) with an optimal Lagrangian multiplier λ_*.

Lemma 4.3 *Let λ_* be an optimal Lagrangian multiplier and $X \in 2^N$ an optimal solution to* (LR(λ_*)). *Then, it holds that*

$$f_1(X) + f_2(X) + \lambda_*(B - c(X)) \geq \textsc{opt}. \tag{36}$$

Moreover, the following properties hold according to the value of $c(X)$:
(i) *if $c(X) < B$, then $f_1(X) + f_2(X) \leq \textsc{opt}$ holds,*
(ii) *if $c(X) = B$, then X satisfies the condition* (30),
(iii) *if $c(X) > B$, then $f_1(X) + f_2(X) \geq \textsc{opt}$ holds.*

Proof. We have (36) since (LR(λ_*)) is a relaxation of (1BM$^\natural$I) and X is an optimal solution to (LR(λ_*)). If $c(X) < B$, then the set X is a feasible solution to (1BM$^\natural$I), and therefore $f_1(X_*) + f_2(X_*) \leq \textsc{opt}$ holds. If $c(X) = B$, then the inequality (36) implies that $f_1(X) + f_2(X) \geq \textsc{opt}$, and therefore the condition (30) holds. If $c(X) > B$, then (36) implies $f_1(X) + f_2(X) \geq \textsc{opt}$ since $\lambda_* \geq 0$. □

4.2 Algorithm. We present an algorithm for computing a set $\tilde{X} \in \mathcal{F}_1 \cap \mathcal{F}_2$ satisfying the condition (30). In the following, we explain each step of the algorithm in detail, and prove the validity of the algorithm as well as the strong polynomiality of the running time.

Step 0: Compute an optimal Lagrangian multiplier λ_* and optimal solutions X_*, Y_* of the problem (LR(λ_*)) with $c(X_*) \leq B \leq c(Y_*)$. If $c(X_*) = B$ then output X_* and stop; if $c(Y_*) = B$ then output Y_* and stop; otherwise, set $X := X_*$ and $Y := Y_*$.

Step 1: Construct an auxiliary graph G_X^Y (definition is given below). Find a zero-length cycle C in G_X^Y with the minimum number of arcs and set $X' := X \oplus C$.

Step 2: If $X' = Y$, then apply an additional patching operation explained in Section 4.3 to obtain a new set $\tilde{X} \in \mathcal{F}_1 \cap \mathcal{F}_2$ satisfying the condition (30). Output $\tilde{X}$ and stop.

Step 3: If $c(X') = B$, then output X' and stop.

Step 4: If $c(X') < B$, then set $X := X'$; if $c(X') > B$, then set $Y := X'$. Go to Step 1.

In Step 0, we compute an optimal Lagrangian multiplier λ_*, which can be done in strongly-polynomial time by Lemma 4.2. We also compute two optimal solutions X_* and Y_* to (LR(λ_*)) satisfying $c(X_*) \leq B \leq c(Y_*)$. This can be done in strongly-polynomial time by solving the problems (LR($\lambda_* + \delta$)) and (LR($\lambda_* - \delta$)). Indeed, if X_* and Y_* are optimal solutions to (LR($\lambda_* + \delta$)) and (LR($\lambda_* - \delta$)), respectively, then Lemma 4.1 and the optimality condition (33) imply that

$$(z_{\rm LR})'_+(\lambda_*) = B - c(X_*) \geq 0, \qquad (z_{\rm LR})'_-(\lambda_*) = B - c(Y_*) \leq 0,$$

i.e., $c(X_*) \leq B \leq c(Y_*)$ holds. If $c(X_*) = B$ (resp., $c(Y_*) = B$), then X_* (resp., Y_*) satisfies the condition (30) by Lemma 4.3 (ii). Otherwise (i.e., $c(X_*) < B < c(Y_*)$), we set $X = X_*$, $Y = Y_*$ and start the loop of Steps 1–4.

We note that at the beginning of the loop, the condition $c(X) < B < c(Y)$ is always satisfied (see the description of Step 4 below). In each iteration of the loop, we repeatedly apply a "patching" operation to increase the value of $c(X)$ (or to decrease $c(Y)$) while keeping the condition that X and Y are optimal solutions to the Lagrangian relaxation problem (LR(λ_*)).

The patching operation is done by using a cycle in an auxiliary graph; given $X, Y \in \mathcal{F}_1 \cap \mathcal{F}_2$, we define an *auxiliary graph* $G_X^Y = (V, A)$ by

$$\begin{aligned}
V &= (X \setminus Y) \cup (Y \setminus X) \cup \{s_a, s_d\}, \\
A &= E_1 \cup E_2 \cup A_1 \cup A_2 \cup D_1 \cup D_2, \\
E_1 &= \{(u, v) \mid u \in X \setminus Y,\ v \in Y \setminus X,\ X - u + v \in \mathcal{F}_1\}, \\
E_2 &= \{(v, u) \mid v \in Y \setminus X,\ u \in X \setminus Y,\ X + v - u \in \mathcal{F}_2\}, \\
A_1 &= \{(s_a, v) \mid v \in Y \setminus X,\ X + v \in \mathcal{F}_1\}, \\
A_2 &= \{(v, s_a) \mid v \in Y \setminus X,\ X + v \in \mathcal{F}_2\}, \\
D_1 &= \{(u, s_d) \mid u \in X \setminus Y,\ X - u \in \mathcal{F}_1\}, \\
D_2 &= \{(s_d, u) \mid u \in X \setminus Y,\ X - u \in \mathcal{F}_2\}.
\end{aligned}$$

where s_a, s_d are new elements not in N. We define the arc length $\omega(a)$ of each arc $a \in A$ by

$$\omega(a) = \begin{cases}
\hat{f}_1(X - u + v) - \hat{f}_1(X) & (a = (u, v) \in E_1), \\
f_2(X + v - u) - f_2(X) & (a = (v, u) \in E_2), \\
\hat{f}_1(X + v) - \hat{f}_1(X) & (a = (s_a, v) \in A_1), \\
f_2(X + v) - f_2(X) & (a = (v, s_a) \in A_2), \\
\hat{f}_1(X - u) - \hat{f}_1(X) & (a = (u, s_d) \in D_1), \\
f_2(X - u) - f_2(X) & (a = (s_d, u) \in D_2),
\end{cases}$$

where the function $\hat{f}_1$ is given by (31). The auxiliary graph defined here is a variant of the one for the valuated matroid intersection problem used in [30] (see also Appendix D). Hence, properties of the auxiliary graph for the valuated matroid intersection problem can be used for the auxiliary graph G_X^Y with some appropriate modification.

A *cycle* in the graph G_X^Y is a directed closed walk which visits each node at most once. In every cycle in G_X^Y, arcs in $E_1 \cup A_1 \cup D_1$ and arcs in $E_2 \cup A_2 \cup D_2$ appear alternately, and therefore every cycle contains an even number of arcs. We call a cycle in G_X^Y *admissible* if the cycle does not visit both of of s_a and s_d. An admissible cycle in G_X^Y with the maximum length with respect to ω is called a *maximum admissible cycle* in G_X^Y.

For an admissible cycle C in G_X^Y, we define an operation $X \oplus C$ ($\subseteq N$) by

$$X \oplus C = X \setminus \{u \in X \setminus Y \mid (u, v) \in C \cap (E_1 \cup D_1)\} \cup \{v \in Y \setminus X \mid (u, v) \in C \cap (E_1 \cup A_1)\}.$$

The following properties are easy to see:

- if C visits neither of s_a and s_d, then $C \subseteq E_1 \cup E_2$ and $|X \oplus C| = |X|$,
- if C visits s_a but not s_d, then $C \subseteq E_1 \cup E_2 \cup A_1 \cup A_2$ and $|X \oplus C| = |X| + 1$,
- if C visits s_d but not s_a, then $C \subseteq E_1 \cup E_2 \cup D_1 \cup D_2$ and $|X \oplus C| = |X| - 1$.

The next property follows from the results in [30, Part I] for the valuated matroid intersection problem.

Lemma 4.4 *Let $X, Y \in \mathcal{F}_1 \cap \mathcal{F}_2$.*
(i) *Let C be a maximum admissible cycle in G_X^Y with the minimum number of arcs, and $\omega(C)$ be the total length of the cycle C. Then, we have*

$$X \oplus C \in \mathcal{F}_1 \cap \mathcal{F}_2, \qquad \hat{f}_1(X \oplus C) + f_2(X \oplus C) = \hat{f}_1(X) + f_2(X) + \omega(C).$$

(ii) *If X is an optimal solution to* (LR(λ_*))*, then there exists no positive-length admissible cycle in G_X^Y.*
(iii) *If Y is an optimal solution to* (LR(λ_*)) *and X is not optimal, then there exists a positive-length admissible cycle in G_X^Y.*

From this lemma we can obtain the following property.

Lemma 4.5 *Let X and Y be two distinct optimal solutions to* (LR(λ_*)). *Then, the length of a maximum admissible cycle in G_X^Y is zero.*

Proof. By Lemma 4.4 (ii), the length of every admissible cycle C in G_X^Y is non-positive, i.e, $\omega(C) \leq 0$. Hence, it suffices to show that there exists an admissible cycle with zero length. We prove this by contradiction.

Assume, to the contrary, that every admissible cycle in G_X^Y has negative length, i.e., $\omega(C) < 0$ for every admissible cycle C. We consider a slight perturbation of the objective function in (LR(λ_*)) so that Y is a unique optimal solution and X is not optimal. This can be done by replacing the function $\hat{f}_1$ with a function $\hat{f}_1^\delta : \mathcal{F}_1 \to \mathbb{R}$ given by

$$\hat{f}_1^\delta(Z) = \hat{f}_1(Z) + \delta|Z \cap Y| - \delta|Z \setminus Y| \qquad (Z \in \mathcal{F}_1),$$

where δ is a sufficiently small positive real number; note that $\hat{f}_1^\delta$ is an M$^\natural$-concave function. By this perturbation the auxiliary graph does not change, whereas the arc length changes; we denote by $\omega^\delta(a)$ $(a \in A)$ the arc length after the perturbation.

By applying Lemma 4.4 (iii) to the perturbed problem, there exists an admissible cycle C in the auxiliary graph G_X^Y which has positive length with respect to ω^δ (i.e., $\omega_\delta(C) > 0$) since Y is optimal and X is not optimal in the perturbed problem. On the other hand, we have

$$\omega^\delta(C) \leq \omega(C) + 2\delta \cdot (|C|/2) = \omega(C) + \delta|C|; \tag{37}$$

this follows from the observation that arcs in $E_1 \cup A_1 \cup D_1$ and arcs in $E_2 \cup A_2 \cup D_2$ appear alternately in C and

$$\omega_\delta(a) \leq \omega(a) + 2\delta \ (\forall a \in E_1 \cup A_1 \cup D_1), \qquad \omega_\delta(a) = \omega(a) \ (\forall a \in E_2 \cup A_2 \cup D_2).$$

In addition, it follows from the inequality $\omega(C) < 0$ and the choice of δ that $\omega(C) + \delta|C| < 0$, which together with (37), implies $\omega_\delta(C) < 0$, a contradiction. □

We show that the patching operation generates a new optimal solution to (LR(λ_*)).

Lemma 4.6 *Let X and Y be two distinct optimal solutions to* (LR(λ_*)), *and C be a zero-length admissible cycle in G_X^Y with the minimum number of arcs. Then, $X \oplus C$ is an optimal solution to* (LR(λ_*)) *such that $X \oplus C \neq X$.*

Proof. The statement follows from Lemma 4.4 (i) and Lemma 4.5. □

We now explain each step of the loop in detail. Recall that X and Y are optimal solutions to the problem (LR(λ_*)) satisfying $c(X) < B < c(Y)$.

In Step 1, we compute a zero-length cycle C in G_X^Y with the minimum number of arcs to obtain a new set $X' = X \oplus C$, which a new optimal solution X' to the problem (LR(λ_*)) by Lemma 4.6. Note that such a cycle C can be computed in strongly-polynomial time by using an appropriate shortest-path algorithm since a zero-length cycle is a maximum cycle by Lemma 4.5.

In Step 2, we check if $X' = Y$ or not. If $X' = Y$, then we apply an additional patching operation to obtain a new set $\tilde{X} \in \mathcal{F}_1 \cap \mathcal{F}_2$ satisfying the condition (30). This additional patching operation updates in strongly-polynomial time the current set X by using the cycle C found in Step 1, in a similar way as in the original patching operation; the difference is that we use only a part of C in the additional patching operation. Details are given in the next section.

In Steps 3 and 4, we compare the value $c(X')$ with B. If $c(X') = B$, then X' satisfies the condition (30) by Lemma 4.3 (ii). Hence, we output X' in such a case. Otherwise, we have either $c(X') < B$ or $c(X') > B$; in the former case we replace X with X' and in the latter case we replace Y with X'. In either case the condition $c(X) < B < c(Y)$ is maintained after the update of X or Y. We note that if $X' \neq Y$, then we have

$$|(X' \setminus Y) \cup (Y \setminus X')| < |(X \setminus Y) \cup (Y \setminus X)|, \qquad |(X \setminus X') \cup (X' \setminus X)| < |(X \setminus Y) \cup (Y \setminus X)|,$$

which implies that the loop of Steps 1–4 are repeated at most n times. Therefore, the algorithm terminates in strongly-polynomial time.

4.3 Additional patching operation. We finally explain the additional patching operation used in the case where $X \oplus C = Y$. In this case, the cycle C contains all nodes in $(X \setminus Y) \cup (Y \setminus X)$. The cycle may contain the node s_a or s_d; in such a case we have $|X| - |Y| = \pm 1$.

Let $a_1, a_2, \ldots, a_{2h} \in A$ be a sequence of arcs in the cycle C, where $2h$ is the number of arcs in C. It is assumed that $a_j \in E_1 \cup A_1 \cup D_1$ if j is odd and $a_j \in E_2 \cup A_2 \cup D_2$ if j is even. For $j = 1, 2, \ldots, h$, let $\alpha_j = \omega(a_{2j-1}) + \omega(a_{2j})$. Since C is a zero-length cycle, we have $\sum_{j=1}^{h} \alpha_j = 0$.

The following property of a sequence of real numbers, known as Gasoline Lemma (cf. [26]), is useful in design and analysis of our patching operation.

LEMMA 4.7 *Let $\alpha_1, \alpha_2, \ldots, \alpha_h \in \mathbb{R}$ be a sequence of real numbers satisfying $\sum_{j=1}^{h} \alpha_j = 0$. Then, there exists some $t \in \{1, 2, \ldots, h\}$ such that*

$$\sum_{j=t}^{t+i} \alpha_{j (\mathrm{mod}\, h)} \geq 0 \qquad (i = 0, 1, \ldots, h-1),$$

where $\alpha_0 = \alpha_h$.

From this lemma, we may assume that

$$\sum_{j=1}^{i} \alpha_j \geq 0 \qquad (\forall i = 1, 2, \ldots, h). \tag{38}$$

In the following, we assume that $C \subseteq E_1 \cup E_2$ for simplicity of the description; the remaining cases can be shown similarly. For $j = 1, 2, \ldots, h$, denote $a_{2j-1} = (u_j, v_j)$ and $a_{2j} = (v_j, u_{j+1})$; note that $a_{2j-1} \in E_1$ and $a_{2j} \in E_2$. Since C contains all nodes in $(X \setminus Y) \cup (Y \setminus X)$, we have

$$X \setminus Y = \{u_1, u_2, \ldots, u_h\}, \qquad Y \setminus X = \{v_1, v_2, \ldots, v_h\};$$

For $j = 1, 2, \ldots, h$, we define $\eta_j \in \mathbb{R}$ by

$$\eta_j = c(v_j) - c(u_j).$$

Then, we have

$$\alpha_j \quad = \quad (f_1(X - u_j + v_j) - f_1(X)) + (f_2(X + v_j - u_{j+1}) - f_2(X)) - \lambda_* \eta_j. \tag{39}$$

Let $t \in \{1, 2, \ldots, h\}$ be the minimum integer such that

$$c(X) + \sum_{j=1}^{t} \eta_j > B. \tag{40}$$

Since

$$c(X) < B < c(Y) = c(X) + \sum_{j=1}^{h} \eta_j,$$

we have $t \geq 1$. In addition, the choice of t implies that

$$c(X) + \sum_{j=1}^{t-1} \eta_j \leq B. \tag{41}$$

We define $\tilde{X}, \tilde{X}_1, \tilde{X}_2 \subseteq N$ by

$$\begin{aligned} \tilde{X} &= X \setminus \{u_1, u_2, \ldots, u_t, u_{t+1}\} \cup \{v_1, \ldots, v_t\}, \\ \tilde{X}_1 &= \tilde{X} \cup \{u_{t+1}\} = X \setminus \{u_1, u_2, \ldots, u_t\} \cup \{v_1, \ldots, v_t\}, \\ \tilde{X}_2 &= \tilde{X} \cup \{u_1\} = X \setminus \{u_2, \ldots, u_t, u_{t+1}\} \cup \{v_1, \ldots, v_t\}. \end{aligned}$$

Note that $\tilde{X} = \tilde{X}_1 \cap \tilde{X}_2$ holds. Putting $C' = \{a_1, a_2, \ldots, a_{2t-1}, a_{2t}\}$, we have

$$\begin{aligned} C' \cap E_1 &= \{a_1, a_3, \ldots, a_{2t-1}\}, & \tilde{X}_1 &= X \setminus \{u \mid (u, v) \in C' \cap E_1\} \cup \{v \mid (u, v) \in C' \cap E_1\}, \\ C' \cap E_2 &= \{a_2, a_4, \ldots, a_{2t}\}, & \tilde{X}_2 &= X \setminus \{u \mid (v, u) \in C' \cap E_2\} \cup \{v \mid (v, u) \in C' \cap E_2\}. \end{aligned}$$

Below we show that the set $\tilde{X}$ satisfies $\tilde{X} \in \mathcal{F}_1 \cap \mathcal{F}_2$ and the condition (30).

Lemma 4.8 *It holds that*

$$\tilde{X}_1 \in \mathcal{F}_1,\ \ \tilde{X}_2 \in \mathcal{F}_2, \tag{42}$$

$$f_1(\tilde{X}_1) = f_1(X) + \sum_{j=1}^{t}(f_1(X - u_j + v_j) - f_1(X)), \tag{43}$$

$$f_2(\tilde{X}_2) = f_2(X) + \sum_{j=1}^{t}(f_2(X - u_{j+1} + v_j) - f_2(X)). \tag{44}$$

Proof. By using the fact that C' is a subpath of a zero-length admissible cycle with the smallest number of arcs, we can show the claims by using a similar proof technique as in [30]. Below we give an outline of the proof for $\tilde{X}_1 \in \mathcal{F}_1$ and the equation (43); proof of $\tilde{X}_2 \in \mathcal{F}_2$ and (44) is similar and omitted.

We consider a subgraph $G'_1 = (V'_1, E'_1)$ of the graph G_X^Y such that

$$\begin{aligned} V'_1 &= \{u_1, u_2, \ldots, u_t\} \cup \{v_1, v_2, \ldots, v_t\} = (X_1 \setminus \tilde{X}_1) \cup (\tilde{X}_1 \setminus X_1), \\ E'_1 &= \{(u, v) \mid u, v \in V'_1,\ (u, v) \in E_1\}. \end{aligned}$$

Note that G'_1 is a bipartite graph, and the arc set $C' \cap E_1 = \{(u_j, v_j) \mid j = 1, 2, \ldots, t\}$ is a perfect matching of G'_1. It can be shown by using the fact that C' is a subpath of a maximum admissible cycle that $C' \cap E_1$ is a maximum-length matching in G'_1. Moreover, we can show that $C' \cap E_1$ is a *unique* maximum-length matching in G'_1; this follows from the fact that C' is a subpath of a maximum admissible cycle with the *smallest number of arcs* (cf. [30, Part II, Sec. 2.1]). By using this fact, we can prove, as in the "unique-max lemma" in [30], that

$$\tilde{X}_1 = X \setminus \{u \mid (u, v) \in C' \cap E_1\} \cup \{v \mid (u, v) \in C' \cap E_1\} \in \mathcal{F}_1,$$

$$f_1(\tilde{X}_1) = f_1(X) + \sum_{j=1}^{t} \omega(u_j, v_j) = f_1(X) + \sum_{j=1}^{t}(f_1(X - u_j + v_j) - f_1(X)).$$

That is, we have $\tilde{X}_1 \in \mathcal{F}_1$ and (43). □

Since $\mathcal{F}_1$ and $\mathcal{F}_2$ are the families of matroid independent sets and $\tilde{X}$ is a common subset of $\tilde{X}_1$ and $\tilde{X}_2$, we have $\tilde{X} \in \mathcal{F}_1 \cap \mathcal{F}_2$ by (42).

We then prove the first inequality in the condition (30). It holds that

$$\begin{aligned} \text{OPT} &\le f_1(X) + f_2(X) + \lambda_*(B - c(X)) + \sum_{j=1}^{t} \alpha_j \\ &= \left[f_1(X) + \sum_{j=1}^{t}(f_1(X - u_j + v_j) - f_1(X))\right] \\ &\quad + \left[f_2(X) + \sum_{j=1}^{t}(f_2(X - u_{j+1} + v_j) - f_2(X))\right] + \lambda_*\left[(B - c(X)) - \sum_{j=1}^{t}\eta_j\right] \\ &= f_1(\tilde{X}_1) + f_2(\tilde{X}_2) + \lambda_*\left[(B - c(X)) - \sum_{j=1}^{t}\eta_j\right] \\ &< f_1(\tilde{X}_1) + f_2(\tilde{X}_2), \end{aligned} \tag{45}$$

where the first inequality is by (36) in Lemma 4.3 and (38), the first equality is by (39), the second equality is by (43) and (44), and the last inequality is by (40). Since $\tilde{X}_1 = \tilde{X} \cup \{u_{t+1}\}$ and $\tilde{X}_2 = \tilde{X} \cup \{u_1\}$, the submodularity of f_1 and f_2 (see Theorem 2.2) implies that

$$\begin{aligned} f_1(\tilde{X}) + f_2(\tilde{X}) &= f_1(\tilde{X}_1) - (f_1(\tilde{X}_1) - f_1(\tilde{X})) + f_2(\tilde{X}_2) - (f_2(\tilde{X}_2) - f_2(\tilde{X})) \\ &\ge f_1(\tilde{X}_1) - (f_1(\{u_{t+1}\}) - f_1(\emptyset)) + f_2(\tilde{X}_2) - (f_2(\{u_1\}) - f_2(\emptyset)) \\ &\ge f_1(\tilde{X}_1) + f_2(\tilde{X}_2) - 2 \cdot \max_{v \in N}(f_1(\{v\}) + f_2(\{v\})) \\ &\ge \text{OPT} - 2 \cdot \max_{v \in N}(f_1(\{v\}) + f_2(\{v\})), \end{aligned}$$

where the last inequality is by (45). Hence, the set $\tilde{X}$ satisfies the first inequality in (30).

Finally, we have

$$c(\tilde{X}) = c(X) + \sum_{j=1}^{t} \eta_j - c(u_{t+1}) \le \left(c(X) + \sum_{j=1}^{t-1} \eta_j \right) + c(v_t) \le B + \max_{v \in N} c(v),$$

where the second inequality is by (41). Hence, $\tilde{X}$ satisfies the second inequality in (30). This concludes the proof of Theorem 1.6.

Acknowledgments. An extended abstract of this paper appeared in Proceedings of the 19th Annual European Symposium on Algorithms (ESA 2011), Lecture Notes in Computer Science 6942, Springer 2011. The author thanks Satoru Iwata and Kazuo Murota for discussions and valuable comments on the preliminary version of the manuscript, and the two anonymous referees for their valuable comments on the presentation of the manuscript. This work is supported by Grant-in-Aid of the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

[1] Ausubel, L. M., P. Milgrom. 2002. Ascending auctions with package bidding. *Front. Theor. Econ.* **1** Article 1.

[2] Berger, A., V. Bonifaci, F. Grandoni, G. Schäfer. 2011. Budgeted matching and budgeted matroid intersection via the gasoline puzzle. *Math. Programming* **128** 355–372.

[3] Bertelsen, A. 2005. Substitutes valuations and $\mathrm{M}^{\natural}$-concavity. M.Sc. Thesis, The Hebrew University of Jerusalem.

[4] Bing, M., D. Lehmann, P. Milgrom. 2004. Presentation and structure of substitutes valuations. *Proc. 5th ACM Conference on Electronic Commerce (EC)*. ACM, New York, NY, 238–239.

[5] Blumrosen, L., N. Nisan. 2007. Combinatorial auction. N. Nisan, T. Roughgarden, É. Tardos, and V. V. Vazirani, eds. *Algorithmic Game Theory*. Cambridge Univ. Press, New York, NY, 267–299.

[6] Calinescu, G., C. Chekuri, M. Pál, J. Vondrák. 2007. Maximizing a submodular set function subject to a matroid constraint. *Proc. 12th International Conference on Integer Programming and Combinatorial Optimization (IPCO)*, Springer, Berlin, 182–196.

[7] Calinescu, G., C. Chekuri, M. Pál, J. Vondrák. 2011. Maximizing a submodular set function subject to a matroid constraint. *SIAM J. Comput.* **40** 1740–1766.

[8] Chekuri, C., J. Vondrák, R. Zenklusen. 2011. Multi-budgeted matchings and matroid intersection via dependent rounding. *Proc. 22nd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1080–1097.

[9] Chekuri, C., J. Vondrák, R. Zenklusen. 2011. Submodular function maximization via the multilinear relaxation and contention resolution schemes. *Proc. 43rd Annual ACM Symposium on Theory of Computing*, ACM, New York, NY, 783–792.

[10] Cramton, P., Y. Shoham, R. Steinberg. 2006. *Combinatorial Auctions*. MIT Press, Cambridge, MA.

[11] Dress, A.W.M., W. Wenzel. 1982. Valuated matroids. *Adv. Math.* **93** 214–250.

[12] Feige, U. 1998. A threshold of $\ln n$ for approximating set cover. *J. ACM* **45** 634–652.

[13] Feige, U. 2009. On maximizing welfare when utility functions are subadditive. *SIAM J. Comput.* **39** 122–142.

[14] Feldman, M., J. Naor, R. Schwartz. 2011. A unified continuous greedy algorithm for submodular maximization. *Proc. 52th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, IEEE Computer Society, Washington, DC, 580–579.

[15] Frank, A., É. Tardos. 1988. Generalized polymatroids and submodular flows. *Math. Programming* **42** 489–563.

[16] Fujishige, S. 2005. *Submodular Functions and Optimization. 2nd Edition.* Elsevier, Amsterdam.

[17] Fujishige, S., Z. Yang. 2003. A note on Kelso and Crawford's gross substitutes condition. *Math. Oper. Res.* **28** 463–469.

[18] Grandoni, F., R. Zenklusen. 2010. Approximation schemes for multi-budgeted independence systems. *Proc. 18th Annual European Symposium on Algorithms (ESA)*, Springer, Berlin, 536–548.

[19] Grötschel, M., L. Lovász, A. Schrijver. 1993. *Geometric Algorithms and Combinatorial Optimization. 2nd Edition.* Springer, Berlin.

[20] Gul, F., E. Stacchetti. 1999. Walrasian equilibrium with gross substitutes. *J. Econ. Theory* **87** 95–124.

[21] Kelso, A. S., V. P. Crawford. 1982. Job matching, coalition formation and gross substitutes. *Econometrica* **50** 1483–1504.

[22] Kulik, A., H. Shachnai, T.Tamir. 2009. Maximizing submodular set functions subject to multiple linear constraints. *Proc. 20th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 545–554.

[23] Kulik, A., H. Shachnai, T.Tamir. 2013. Approximations for monotone and non-monotone submodular maximization with knapsack constraints. *Math. Oper. Res.*, published online.

[24] Lee, J., V. S. Mirrokni, V. Nagarajan, M. Sviridenko. 2010. Maximizing nonmonotone submodular functions under matroid or knapsack constraints. *SIAM J. Discrete Math.* **23** 2053–2078.

[25] Lehmann, B., D. Lehmann, N. Nisan. 2006. Combinatorial auctions with decreasing marginal utilities. *Games Econom. Behav.* **55** 270–296.

[26] Lin, S., B.W. Kernighan. 1973. An effective heuristic algorithm for the traveling salesman problem. *Oper. Res.* **21** 498–516.

[27] Megiddo, N. 1979. Combinatorial optimization with rational objective functions. *Math. Oper. Res.* **4** 414–424.

[28] Moriguchi, S., N. Tsuchimura. 2009. Discrete L-convex function minimization based on continuous relaxation. *Pacific J. Optim.* **5** 227–236.

[29] Murota, K. 1996. Convexity and Steinitz's exchange property. *Adv. Math.* **124** 272–311.

[30] Murota, K. 1996. Valuated matroid intersection, I: optimality criteria, II: algorithms. *SIAM J. Discrete Math.* **9** 545–561, 562–576.

[31] Murota, K. 1998. Discrete convex analysis. *Math. Programming* **83** 313–371.

[32] Murota, K. 2003. *Discrete Convex Analysis.* Society for Industrial and Applied Mathematics, Philadelphia, PA.

[33] Murota, K. 2009. Recent developments in discrete convex analysis. W. J. Cook, J. Lovász, and J. Vygen, eds. *Research Trends in Combinatorial Optimization.* Springer, Berlin, 219–260.

[34] Murota, K., A. Shioura. 1999. M-convex function on generalized polymatroid. *Math. Oper. Res.* **24** 95–105.

[35] Murota, K., A. Shioura. 2000. Extension of M-convexity and L-convexity to polyhedral convex functions. *Adv. Appl. Math.* **25** 352–427.

[36] Murota, K., A. Tamura. 2003. Application of M-convex submodular flow problem to mathematical economics. *Japan J. Indust. Appl. Math.* **20** 257–277.

[37] Murota, K., A. Tamura. 2003. New characterizations of M-convex functions and their applications to economic equilibrium models with indivisibilities. *Discrete Appl. Math.* **131** 495–512.

[38] Oxley, J. 1992. *Matroid Theory.* Oxford Univ. Press, New York.

[39] Ravi, R., M. X. Goemans. 1996. The constrained minimum spanning tree problem. *Proc. 5th Scandinavian Workshop on Algorithm Theory (SWAT)*, Springer, Berlin, 66–75.

[40] Rockafellar, R.T. 1970. *Convex Analysis.* Princeton Univ. Press, Princeton, NJ.

[41] Schrijver, A. 2003. *Combinatorial Optimization: Polyhedra and Efficiency.* Springer, Berlin.

[42] Shioura, A. 2009. On the pipage rounding algorithm for submodular function maximization: a view from discrete convex analysis. *Discrete Math. Algorithms Appl.* **1** 1–23.

[43] Sviridenko, M. 2004. A note on maximizing a submodular set function subject to a knapsack constraint. *Oper. Res. Lett.* **32** 41–43.

[44] Wolsey, L.A. 1982. Maximising real-valued submodular functions: primal and dual heuristics for location problems. *Math. Oper. Res.* **7** 410–425.

Appendix A. Reduction of budgeted M$^\natural$-concave maximization to budgeted GS utility maximization. We show that the k-budgeted M$^\natural$-concave maximization problem (kBM$^\natural$M) can be reduced to the k-budgeted GS utility maximization problem (1).

Given an instance of (kBM$^\natural$M) with an M$^\natural$-concave function $f : \mathcal{F} \to \mathbb{R}$, let $\tilde{f} : 2^N \to \mathbb{R}$ be a function given by (2), i.e.,

$$\tilde{f}(X) = \max\{f(Y) \mid Y \in \mathcal{F},\ Y \subseteq X\} \qquad (X \in 2^N).$$

It should be noted that the function value of $\tilde{f}$ can be evaluated in polynomial time (see Theorem 2.1) since the value $\tilde{f}(X)$ is given as the maximization of an M$^\natural$-concave function $f^X : \mathcal{F}^X \to \mathbb{R}$ defined by

$$\mathcal{F}^X = \{Y \mid Y \in \mathcal{F},\ Y \subseteq X\}, \qquad f^X(Y) = f(Y)\ (Y \in \mathcal{F}^X).$$

Proposition A.1 *The function $\tilde{f} : 2^N \to \mathbb{R}$ is a GS utility function.*

Proof. By Theorem 1.1, it suffices to show that the function $\tilde{f}$ satisfies the condition (M$^\natural$-EXC) in the definition of M$^\natural$-concave functions. Take $X, Y \in 2^N$ and $u \in X \setminus Y$. Let $I, J \in \mathcal{F}$ be subsets of X and Y, respectively, such that $\tilde{f}(X) = f(I)$ and $\tilde{f}(Y) = f(J)$.

If $u \notin I$, then

$$\tilde{f}(X - u) \geq f(I) = \tilde{f}(X), \quad \tilde{f}(Y + u) \geq f(J) = \tilde{f}(Y),$$

which implies (i) in (M$^\natural$-EXC).

We then assume $u \in I$. Since $u \in X \setminus Y$, we have $u \in I \setminus J$. By (M$^\natural$-EXC) applied to f, I, J, and u, at least one of (a) and (b) holds, where

(a) $I - u \in \mathcal{F},\ J + u \in \mathcal{F}$, and $f(I) + f(J) \leq f(I - u) + f(J + u)$,
(b) $\exists v \in J \setminus I$: $I - u + v \in \mathcal{F}$, $J + u - v \in \mathcal{F}$, and $f(I) + f(J) \leq f(I - u + v) + f(J + u - v)$.

If (a) holds, then we have

$$\tilde{f}(X - u) + \tilde{f}(Y + u) \geq f(I - u) + f(J + u) \geq f(I) + f(J) = \tilde{f}(X) + \tilde{f}(Y),$$

i.e., (i) in (M$^\natural$-EXC) holds.

We then consider the case where (b) holds. If $v \in X$, then $I - u + v \subseteq X - u$, $J + u - v \subseteq Y + u$, and hence

$$\tilde{f}(X - u) \geq f(I - u + v), \qquad \tilde{f}(Y + u) \geq f(J + u - v),$$

which implies

$$\tilde{f}(X - u) + \tilde{f}(Y + u) \geq f(I - u + v) + f(J + u - v) \geq f(I) + f(J) = \tilde{f}(X) + \tilde{f}(Y),$$

i.e., (i) in (M$^\natural$-EXC) holds. If $v \notin X$, then we have $v \in Y \setminus X$ and

$$\tilde{f}(X - u + v) \geq f(I - u + v), \qquad \tilde{f}(Y + u - v) \geq f(J + u - v),$$

which implies

$$\tilde{f}(X - u + v) + \tilde{f}(Y + u - v) \geq f(I - u + v) + f(J + u - v) \geq f(I) + f(J) = \tilde{f}(X) + \tilde{f}(Y),$$

i.e., (ii) in (M$^\natural$-EXC) holds. □

We consider the k-budgeted GS utility maximization problem (1) with the objective function $\tilde{f}$:

$$\text{Maximize } \tilde{f}(X) \quad \text{subject to } X \in 2^N,\ c_i(X) \leq B_i\ (i = 1, 2, \ldots, k). \tag{46}$$

The following property shows that an optimal solution to (kBM$^\natural$M) can be obtained by solving the problem (46).

Proposition A.2 *Every minimal optimal solution to the problem* (46) *is an optimal solution to* (kBM$^\natural$M).

Proof. Let $X_* \in 2^N$ be a minimal optimal solution to the problem (46). Since $f(Y) \leq \tilde{f}(Y)$ for every $Y \in \mathcal{F}$, it suffices to show that $X_* \in \mathcal{F}$ and $\tilde{f}(X_*) = f(X_*)$. Let $Y_* \in \mathcal{F}$ be a subset of X_* such that $\tilde{f}(X_*) = f(Y_*)$. We have

$$c_i(Y_*) \leq c_i(X_*) \leq B_i \qquad (i = 1, 2, \ldots, k),$$

and it holds that $\tilde{f}(Y_*) = f(Y_*)$ by the definition of $\tilde{f}$. Hence, the set Y_* is also an optimal solution to (46). By the minimality of X_*, we have $X_* = Y_* \in \mathcal{F}$, which implies $\tilde{f}(X_*) = f(X_*)$. □

Appendix B. Partial enumeration technique for PTAS. Theorems 1.2 and 1.6 state that there exist polynomial-time algorithms which compute high-quality solutions which are almost feasible to (kBM$^\natural$M) and to (1BM$^\natural$I), respectively. We show that by using a standard technique called "partial enumeration" (see, e.g., [2, 18, 39]), these algorithms can be transformed into PTASes for (kBM$^\natural$M) and for (1BM$^\natural$I), respectively.

We here consider a more general setting. Let $\mathcal{F} \subseteq 2^N$ be an *independence system*, i.e., $\mathcal{F}$ satisfies the condition that if $X \in \mathcal{F}$ and $Y \subseteq X$ then $Y \in \mathcal{F}$. Let $f : \mathcal{F} \to \mathbb{R}$ be a function defined on $\mathcal{F}$ satisfying $f(\emptyset) = 0$, and suppose that f is a submodular function in the following sense:

$$f(X) + f(Y) \geq f(X \cap Y) + f(X \cup Y) \quad (\forall X, Y \in \mathcal{F} \text{ such that } X \cup Y \in \mathcal{F}). \tag{47}$$

For $X \in \mathcal{F}$ and $Y \subseteq N$ with $X \subseteq Y$, we define a set family $\mathcal{F}_X^Y$ $(\subseteq 2^{Y \setminus X})$ and a function $f_X^Y : \mathcal{F}_X^Y \to \mathbb{R}$ by

$$\mathcal{F}_X^Y = \{U \mid U \subseteq Y \setminus X,\ U \cup X \in \mathcal{F}\}, \tag{48}$$

$$f_X^Y(U) = f(X \cup U) - f(X) \quad (U \in \mathcal{F}_X^Y). \tag{49}$$

Note that $\mathcal{F}_X^Y$ is also an independence system and f_X^Y is a submodular function on $\mathcal{F}_X^Y$ with $f_X^Y(\emptyset) = 0$. We say that $\mathcal{F}_X^Y$ (resp., f_X^Y) is a *minor* of $\mathcal{F}$ (resp., f).

Let $\mathcal{S}$ be a family of submodular functions $f : \mathcal{F} \to \mathbb{R}$ defined on independence systems $\mathcal{F}$ such that $f(\emptyset) = 0$, and assume that $\mathcal{S}$ is *minor-closed*, i.e., every minor of $f \in \mathcal{S}$ is also in $\mathcal{S}$. We consider the following budgeted optimization problem:

$$(k\text{BSM}) \text{ Maximize } \ f(X) \quad \text{ subject to } \ X \in \mathcal{F},\ c_i(X) \leq B_i \ (1 \leq i \leq k), \tag{50}$$

where $f : \mathcal{F} \to \mathbb{R}$ is a function in $\mathcal{S}$, k is a positive integer, and $c_i \in \mathbb{R}_+^N$, $B_i \in \mathbb{R}_+$ $(i = 1, 2, \ldots, k)$. We denote by $\textsc{opt}$ the optimal value of (kBSM). Note that the problems (kBM$^\natural$M) and (1BM$^\natural$I) are special cases of (kBSM). We may assume that

$$\{v\} \text{ is a feasible solution to } (k\text{BSM}) \text{ such that } f(\{v\}) > 0 \qquad (\forall v \in N);$$

the validity of this assumption can be shown in a similar way as in Proposition 3.1.

We prove the following theorem by applying the partial enumeration technique to (kBSM). We define two parameters Φ and Ψ representing the input size of the problem by

$$\Phi = \max_{X \in \mathcal{F}} \langle f(X) \rangle, \qquad \Psi = \max \left[\max_{1 \leq i \leq k, j \in N} \langle c_i(j) \rangle,\ \max_{1 \leq i \leq k} \langle B_i \rangle \right].$$

Theorem B.1 *Let $\alpha \in [0,1]$ and $\eta \in \mathbb{Z}_+$. Suppose that the problem* (kBSM) *has an algorithm which computes a set $\tilde{X} \in \mathcal{F}$ satisfying*

$$f(\tilde{X}) \geq \alpha \cdot \textsc{opt} - \eta \cdot \max_{v \in N} f(\{v\}),$$

$$c_i(\tilde{X}) \leq B_i + \eta \cdot \max_{v \in N} c_i(v) \qquad (i = 1, 2, \ldots, k)$$

in $\mathrm{O}(\mu(n, \Phi, \Psi))$ time, where $\mu(n, \Phi, \Psi)$ is a function which is monotone nondecreasing with respect to n, Φ, and Ψ.
(i) *For every $\varepsilon \in (0, \alpha]$, the problem* ($k$BSM) *has an $(\alpha - \varepsilon)$-approximation algorithm which runs in $n^{\mathrm{O}(k/\varepsilon)} \cdot \mathrm{O}(\mu(n, \Phi, \Psi))$ time.*
(ii) *If $\mu(n, \Phi, \Psi)$ is a polynomial function in n, Φ, and Ψ, then* (kBSM) *has a polynomial-time $(\alpha - \varepsilon)$-approximation algorithm for fixed k and ε.*

Then, Theorem 1.3 (resp., Theorem 1.5) is an immediate consequence of Theorem B.1 and Theorem 1.2 (resp., Theorem 1.6), where $\alpha = 1 - \varepsilon$ and $\eta = 2k$ (resp., $\eta = 2$).

We now give a proof of Theorem B.1. We set

$$\varepsilon' = \frac{1}{\lceil (\alpha + 1)/\varepsilon \rceil}$$

so that $1/\varepsilon' = \lceil (\alpha + 1)/\varepsilon \rceil$ is a positive integer. Let $X_* \in \mathcal{F}$ be an optimal solution of (kBSM) which is fixed in the following discussion. We may assume that $|X_*| > (k+1)\eta/\varepsilon'$ since otherwise the cardinality of X_* is bounded by a constant number and therefore such X_* can be found by a brute-force algorithm in polynomial time.

Our algorithm consists of the following three major steps:

Step 1: Guess a subset X_b of X_* with $|X_b| = (k+1)\eta/\varepsilon'$ consisting of "large" elements. Intuitively, X_b consists of elements $v \in N$ such that at least one of the the values $f(\{v\})$ and $c_i(v)$ $(i = 1, 2, \ldots, k)$ is sufficiently large compared to other elements in N (a precise definition of X_b is given later).

Step 2: By using the algorithm in the assumption of Theorem B.1, compute a set X_s satisfying the following conditions:

$$X_b \cup X_s \in \mathcal{F}, \tag{51}$$
$$f(X_b \cup X_s) \geq (\alpha - \varepsilon')\textsc{opt}, \tag{52}$$
$$c_i(X_b \cup X_s) \leq (1 + \varepsilon')B_i \qquad (i = 1, 2, \ldots, k). \tag{53}$$

Note that the set $X_b \cup X_s$ may violate the constraints of (kBSM) (but only slightly). If $X_b \cup X_s$ is a feasible solution to (kBSM), then output $X_b \cup X_s$ and stop; otherwise, go to Step 3.

Step 3: To make the set $X_b \cup X_s$ a feasible solution to the problem (kBSM), compute a subset U of $X_b \cup X_s$ such that $(X_b \cup X_s) \setminus U$ is an $(1 - \varepsilon')(\alpha - \varepsilon')$-approximate feasible solution to (kBSM). Output $(X_b \cup X_s) \setminus U$.

If the set X_b is guessed correctly, then the output in Step 3 is an $(\alpha - \varepsilon)$-approximate solution since $(1 - \varepsilon')(\alpha - \varepsilon') \geq \alpha - \varepsilon$.

It should be noted that for a given set X_b, it is difficult to check if X_b is a correct guess. Hence, we need to enumerate all possible subsets X_b of N with cardinality $(k+1)\eta/\varepsilon'$, and for each subset X_b we apply Steps 2 and 3 to obtain a feasible solution to (kBSM). That is, we obtain at most $n^{(k+1)\eta/\varepsilon'} = n^{\mathrm{O}(k/\varepsilon)}$ feasible solutions, and at least one of them is an $(\alpha - \varepsilon)$-approximate solution to (kBSM). Therefore, we just need to output the best feasible solution among the feasible solutions obtained so far.

Below we explain the details of each step.

Details of Step 1. We explain how to compute a set X_b in Step 1.

As a part of the set X_b, we first guess a subset Z_0 of X_* which maximizes the value $f(Z_0)$ under the condition that

$$|Z_0| = \eta/\varepsilon', \qquad Z_0 \text{ is a feasible solution to } (k\text{BSM}).$$

This is done by enumerating all subsets of N with cardinality η/ε'.

Let

$$N_0 = \{v \in N \setminus Z_0 \mid Z_0 \cup \{v\} \in \mathcal{F},\ f(Z_0 \cup \{v\}) - f(Z_0) \leq (\varepsilon'/\eta)f(Z_0)\}.$$

Lemma B.1 *$X_* \setminus Z_0 \subseteq N_0$ holds if Z_0 is guessed correctly.*

Proof. Assume, to the contrary, that there exists some $v \in X_* \setminus Z_0$ such that $v \notin N_0$. We have $Z_0 \cup \{v\} \in \mathcal{F}$ since it is a subset of $X_* \in \mathcal{F}$. Since $v \notin N_0$, we have

$$f(Z_0 \cup \{v\}) - f(Z_0) > \frac{\varepsilon'}{\eta} f(Z_0). \tag{54}$$

Let $u = u_* \in Z_0$ minimize the value $f(Z_0) - f(Z_0 \setminus \{u\})$. It follows from the submodularity (47) that

$$\begin{aligned} f(Z_0) - f(Z_0 \setminus \{u_*\}) \quad &\leq \quad \frac{1}{|Z_0|} \sum_{u \in Z_0} (f(Z_0) - f(Z_0 \setminus \{u\})) \\ &\leq \quad \frac{1}{|Z_0|}(f(Z_0) - f(\emptyset)) \ = \ \frac{\varepsilon'}{\eta} f(Z_0). \end{aligned} \tag{55}$$

From (54) and (55) follows that

$$f(Z_0) - f(Z_0 \setminus \{u_*\}) < f(Z_0 \cup \{v\}) - f(Z_0) \leq f((Z_0 \setminus \{u_*\}) \cup \{v\}) - f(Z_0 \setminus \{u_*\}),$$

where the last inequality is by submodularity (47). Hence, we have $f(Z_0) < f((Z_0 \setminus \{u_*\}) \cup \{v\})$, a contradiction to the choice of Z_0 since $(Z_0 \setminus \{u_*\}) \cup \{v\}$ is a feasible solution to to (kBSM) with cardinality equal to η/ε'. □

Based on the lemma above, we select remaining elements of X_b from the set N_0. We then guess a set Z_1 of η/ε' largest elements in $X_* \setminus Z_0$ with respect to the cost c_1. This is done by selecting a subset Z_1 of N_0 satisfying

$$|Z_1| = \eta/\varepsilon', \qquad Z_0 \cup Z_1 \text{ is a feasible solution to } (k\text{BSM}).$$

Let

$$N_1 = \{v \in N_0 \setminus Z_1 \mid Z_0 \cup Z_1 \cup \{v\} \in \mathcal{F},\ c_1(v) \leq \min_{u \in Z_1} c_1(u)\}.$$

If Z_1 is a correct guess, then we have $X_* \setminus (Z_0 \cup Z_1) \subseteq N_1$ since Z_1 is chosen as the set of largest elements in $X_* \setminus Z_0$ with respect to the cost c_1.

In a similar way, we iteratively guess a set Z_i of η/ε' largest elements in $X_* \setminus (Z_0 \cup Z_1 \cup \cdots \cup Z_{i-1})$ with respect to the cost c_i for $i = 2, 3, \ldots, k$. This is done by selecting a subset Z_i of N_{i-1} satisfying

$$|Z_i| = \eta/\varepsilon', \qquad Z_0 \cup Z_1 \cup \cdots \cup Z_i \text{ is a feasible solution to } (k\text{BSM}).$$

Let

$$N_i = \{v \in N_{i-1} \setminus Z_i \mid Z_0 \cup Z_1 \cup \cdots \cup Z_i \cup \{v\} \in \mathcal{F},\ c_i(v) \leq \min_{u \in Z_i} c_i(u)\}.$$

If Z_i is a correct guess, then we have

$$X_* \setminus (Z_0 \cup Z_1 \cup \cdots \cup Z_{i-1} \cup Z_i) \subseteq N_i.$$

Let

$$X_b = \bigcup_{i=0}^{k} Z_i.$$

Due to the choice of $Z_0, Z_1, \ldots, Z_k$, we see that X_b is a feasible solution to (kBSM), even if X_b is not a correct guess. If $X_b \subseteq X_*$, then we have

$$f(X_b) \geq f(X_*) - f(X_* \setminus X_b) \geq 0, \tag{56}$$

where the first inequality is by the submodularity of f and the second by the optimality of X_*.

Details of Step 2. We then explain how to compute a set X_s in Step 2. We denote

$$\mathcal{F}' = \mathcal{F}_{X_b}^{X_b \cup N_k}, \qquad f' = f_{X_b}^{X_b \cup N_k}$$

(see (48) and (49) for the definitions of $\mathcal{F}_{X_b}^{X_b \cup N_k}$ and $f_{X_b}^{X_b \cup N_k}$). Then, f' is a function defined on $\mathcal{F}'$ and satisfies $f' \in \mathcal{S}$. We consider an instance of (kBSM) given by

$$\text{Maximize } \ f'(U) \quad \text{ subject to } \ U \in \mathcal{F}',\ c_i(U) \leq B'_i\ (1 \leq i \leq k),$$

where $B'_i = B_i - c_i(X_b)$ for each i. We denote by OPT' the optimal value of this instance. Then, $\text{OPT}' + f(X_b) = \text{OPT}$ holds, provided that the set X_b is guessed correctly.

The assumption of Theorem B.1 implies that we can compute in $\mathrm{O}(\mu(n, \Phi, \Psi))$ time a set $X_s \in \mathcal{F}'$ satisfying

$$f'(X_s) \geq \alpha \cdot \text{OPT}' - \eta \cdot \max_{v \in N_k} f'(\{v\}), \tag{57}$$

$$c_i(X_s) \leq B'_i + \eta \cdot \max_{v \in N_k} c_i(v) \qquad (i = 1, 2, \ldots, k). \tag{58}$$

We show that this set X_s satisfies the conditions (51), (52), and (53) if the set X_b is guessed correctly.

Since $X_s \in \mathcal{F}'$, we have $X_b \cup X_s \in \mathcal{F}$, i.e., (51) holds. We have

$$\begin{aligned} \max_{v \in N_k} f'(\{v\}) &= \max_{v \in N_k} \{f(X_b \cup \{v\}) - f(X_b)\} \leq \max_{v \in N_k} \{f(Z_0 \cup \{v\}) - f(Z_0)\} \\ &\leq \max_{v \in N_0} \{f(Z_0 \cup \{v\}) - f(Z_0)\} \leq \frac{\varepsilon'}{\eta} f(Z_0), \end{aligned} \tag{59}$$

where the first inequality is by the submodularity of f and $Z_0 \subseteq X_b$, the second by $N_k \subseteq N_0$, and the last by the definition of N_0. Hence, (52) can be shown as follows:

$$\begin{aligned} f(X_b \cup X_s) &= f'(X_s) + f(X_b) \\ &\geq \alpha \cdot \text{OPT}' - \eta \cdot \max_{v \in N_k} f'(\{v\}) + f(X_b) \\ &= \alpha \cdot \text{OPT} + (1-\alpha) f(X_b) - \eta \cdot \max_{v \in N_k} f'(\{v\}) \\ &\geq \alpha \cdot \text{OPT} - \eta \cdot \frac{\varepsilon'}{\eta} f(Z_0) = \alpha \cdot \text{OPT} - \varepsilon' f(Z_0) \geq (\alpha - \varepsilon')\text{OPT}, \end{aligned}$$

where the first inequality is by (57), the second by (56) and (59), and the last by $f(Z_0) \leq \textsc{opt}$.

For $i = 1, 2, \ldots, k$, we have

$$\max_{v \in N_k} c_i(v) \leq \max_{v \in N_i} c_i(v) \leq \min_{u \in Z_i} c_i(u) \leq \frac{1}{|Z_i|} c_i(Z_i) = \frac{\varepsilon'}{\eta} c_i(Z_i) \tag{60}$$

by the choice of Z_i and $N_k \subseteq N_i$. It follows from (58), (60), and $c_i(Z_i) \leq B_i$ that

$$\begin{aligned} c_i(X_b \cup X_s) \ = \ c_i(X_b) + c_i(X_s) \ \leq \ c_i(X_b) + B'_i + \eta \cdot \max_{v \in N_k} c_i(v) \ &= \ B_i + \eta \cdot \max_{v \in N_k} c_i(v) \\ &\leq \ B_i + \varepsilon' c_i(Z_i) \leq (1 + \varepsilon') B_i. \end{aligned}$$

That is, (53) holds.

Details of Step 3. Suppose that $X_b \cup X_s$ is not a feasible solution to (kBSM). In Step 3, we finally construct an $(1 - \varepsilon')(\alpha - \varepsilon')$-approximate feasible solution by deleting some elements in $X_b \cup X_s$. Let $\{U_1, U_2, \ldots, U_{(1/\varepsilon')-1}, U_{1/\varepsilon'}\}$ be an arbitrarily chosen partition of X_b such that $|U_j \cap Z_h| = \eta$ for each j and h; recall that $|Z_h| = \eta/\varepsilon'$ for all $h = 0, 1, \ldots, k$ and therefore such a partition exists. We also set $t = (1/\varepsilon') + 1$ and $U_t = X_s$. Then, $\{U_1, U_2, \ldots, U_t\}$ is a partition of $X_b \cup X_s$.

For each $j = 1, 2, \ldots, t$, we have $(X_b \cup X_s) \setminus U_j \in \mathcal{F}$ since $X_b \cup X_s \in \mathcal{F}$. To conclude the proof of Theorem B.1, it suffices to show that the following inequalities hold:

$$c_i((X_b \cup X_s) \setminus U_j) \leq B_i \qquad (\forall i = 1, 2, \ldots, k,\ \forall j = 1, 2, \ldots, t), \tag{61}$$

$$\max_{1 \leq j \leq t} f((X_b \cup X_s) \setminus U_j) \geq (1 - \varepsilon')(\alpha - \varepsilon')\textsc{opt}. \tag{62}$$

The inequality (61) with $j = t$ follows immediately from the fact that $(X_b \cup X_s) \setminus U_t = X_b$ is a feasible solution to (kBSM). For each $i = 1, 2, \ldots, k$ and $j = 1, 2, \ldots, 1/\varepsilon'$, it holds that

$$c_i(U_j) \geq c_i(U_j \cap Z_i) \geq \eta \cdot \min_{u \in Z_i} c_i(u) \geq \eta \cdot \max_{v \in N_i} c_i(v) \geq \eta \cdot \max_{v \in N_k} c_i(v),$$

which, together with (58), implies that

$$c_i((X_b \cup X_s) \setminus U_j) = c_i(X_b \cup X_s) - c_i(U_j) \leq B_i + \eta \cdot \max_{v \in N_k} c_i(v) - \eta \cdot \max_{v \in N_k} c_i(v) = B_i.$$

Hence, the inequality (61) holds.

The inequality (62) can be shown by using the following property of f:

Lemma B.2 (cf. [13]) *Let $f : \mathcal{F} \to \mathbb{R}$ be a submodular function defined on an independence system $\mathcal{F} \subseteq 2^N$ in the sense of* (47). *Also, let $U, V_1, V_2, \ldots, V_t$ be subsets of N, and $\lambda_1, \lambda_2, \ldots, \lambda_t$ be nonnegative real numbers such that $\sum_{j=1}^{t} \lambda_j = 1$ and $\sum_{j=1}^{t} \lambda_j \chi_{V_j} = \chi_U$. Then, it holds that*

$$f(U) \leq \sum_{j=1}^{t} \lambda_j f(V_j).$$

We can obtain the inequality (62) as follows:

$$\begin{aligned} \max_{1 \leq j \leq t} f((X_b \cup X_s) \setminus U_j) \ \geq \ \frac{1}{t} \sum_{j=1}^{t} f((X_b \cup X_s) \setminus U_j) \ &\geq \ \frac{1}{t} \cdot (t-1) f(X_b \cup X_s) \\ &\geq \ (1 - \varepsilon') f(X_b \cup X_s) \ \geq \ (1 - \varepsilon')(\alpha - \varepsilon')\textsc{opt}, \end{aligned}$$

where the second inequality is by Lemma B.2 and the last inequality by (52).

Appendix C. Equivalence between M$^\natural$-concave function and valuated matroid. We give a rigorous proof for the equivalence between M$^\natural$-concave function and valuated matroid by showing that every M$^\natural$-concave function defined on a family of matroid independent sets can be transformed to a valuated matroid which has the same information, and vice versa.

From M$^\natural$-concave function to valuated matroid. Let $f : \mathcal{F} \to \mathbb{R}$ be an M$^\natural$-concave function defined on matroid independent sets $\mathcal{F}$. We define a valuated matroid $g : \mathcal{B} \to \mathbb{R}$ having the same information as f in the following way.

Let $k = \max\{|X| \mid X \in \mathcal{F}\}$. Also, let $s_1, s_2, \ldots, s_k$ be elements not in N, $S = \{s_1, s_2, \ldots, s_k\}$, and $\tilde{N} = N \cup S$. Define $\mathcal{B} \subseteq 2^{\tilde{N}}$ and a function $g : \mathcal{B} \to \mathbb{R}$ by

$$\begin{aligned} \mathcal{B} &= \{\tilde{X} \subseteq \tilde{N} \mid |\tilde{X}| = k,\ \tilde{X} \cap N \in \mathcal{F}\}, & (63)\\ g(\tilde{X}) &= f(\tilde{X} \cap N) \qquad (\tilde{X} \in \mathcal{B}). & (64)\end{aligned}$$

We show that $\mathcal{B}$ is a base family of some matroid and g is a valuated matroid. The proof below is based on the following property of M$^\natural$-concave functions.

LEMMA C.1 ([34]) *Let $f : \mathcal{F} \to \mathbb{R}$ be an M$^\natural$-concave function defined on matroid independent sets $\mathcal{F}$, and $X, Y \in \mathcal{F}$.*
(i) *If $|X| \le |Y|$, then for every $u \in X \setminus Y$ there exists some $v \in Y \setminus X$ such that*

$$X - u + v \in \mathcal{F},\ Y + u - v \in \mathcal{F},\ \text{and } f(X) + f(Y) \le f(X - u + v) + f(Y + u - v).$$

(ii) *If $|X| < |Y|$, then there exists some $v \in Y \setminus X$ such that*

$$X + v \in \mathcal{F},\ Y - v \in \mathcal{F},\ \text{and } f(X) + f(Y) \le f(X + v) + f(Y - v).$$

Let $\tilde{X}, \tilde{Y} \in \mathcal{B}$ and $u \in \tilde{X} \setminus \tilde{Y}$. It suffices to prove that the following condition holds:

$$\exists v \in \tilde{Y} \setminus \tilde{X} \text{ such that } \tilde{X} - u + v \in \mathcal{B},\ \tilde{Y} + u - v \in \mathcal{B},\ g(\tilde{X}) + g(\tilde{Y}) \le g(\tilde{X} - u + v) + g(\tilde{Y} + u - v).$$

We show below the following equivalent condition in terms of $\mathcal{F}$ and f:

$$\begin{aligned} &\exists v \in \tilde{Y} \setminus \tilde{X} \text{ such that } (\tilde{X} - u + v) \cap N \in \mathcal{F}, \quad (\tilde{Y} + u - v) \cap N \in \mathcal{F},\\ &\qquad f(X) + f(Y) \le f((\tilde{X} - u + v) \cap N) + f((\tilde{Y} + u - v) \cap N), \end{aligned} \tag{65}$$

where $X = \tilde{X} \cap N$ and $Y = \tilde{Y} \cap N$. By definition, we have $X, Y \in \mathcal{F}$.

[Case 1: $u \in N$] We have $u \in X \setminus Y$. By (M$^\natural$-EXC) applied to f, we have either (a) or (b) (or both) holds:

(a) $X - u \in \mathcal{F},\ Y + u \in \mathcal{F}$, and $f(X) + f(Y) \le f(X - u) + f(Y + u)$,
(b) $\exists v \in Y \setminus X$: $X - u + v \in \mathcal{F}, Y + u - v \in \mathcal{F}$, and $f(X) + f(Y) \le f(X - u + v) + f(Y + u - v)$.

By Lemma C.1 (i), the statement (b) always holds whenever $|X| \le |Y|$.

Suppose that (a) occurs. Then, we may assume $|X| > |Y|$. Since $|\tilde{X}| = |\tilde{Y}|$, there exists some $v = s_h \in (\tilde{Y} \setminus \tilde{X}) \cap S$. With this v we have

$$(\tilde{X} - u + v) \cap N = X - u \in \mathcal{F}, \qquad (\tilde{Y} + u - v) \cap N = Y + u \in \mathcal{F}.$$

Since $f(X) + f(Y) \le f(X - u) + f(Y + u)$ holds by assumption, we have (65).

We then suppose that (b) occurs. The element $v \in Y \setminus X$ in (b) satisfies $v \in \tilde{Y} \setminus \tilde{X}$, and

$$\begin{aligned} &(\tilde{X} - u + v) \cap N = X - u + v \in \mathcal{F}, \qquad (\tilde{Y} + u - v) \cap N = Y + u - v \in \mathcal{F},\\ &f(X) + f(Y) \le f(X - u + v) + f(Y + u - v). \end{aligned}$$

Hence, (65) holds as well.

[Case 2: $u \in S$] Suppose that there exists some $v \in (\tilde{Y} \setminus \tilde{X}) \cap S$. Then, we have (65) since

$$(\tilde{X} - u + v) \cap N = X \in \mathcal{F}, \qquad (\tilde{Y} + u - v) \cap N = Y \in \mathcal{F}.$$

Suppose that $(\tilde{Y} \setminus \tilde{X}) \cap S = \emptyset$. We have $\tilde{Y} \cap S \subseteq (\tilde{X} \cap S) \setminus \{u\}$, implying that $|\tilde{Y} \cap S| < |\tilde{X} \cap S|$. Since $|\tilde{X}| = |\tilde{Y}|$, it holds that

$$|X| = |\tilde{X}| - |\tilde{X} \cap S| < |\tilde{Y}| - |\tilde{Y} \cap S| = |Y|.$$

By Lemma C.1 (ii), there exists some $v \in Y \setminus X$ such that

$$X + v \in \mathcal{F}, \quad Y - v \in \mathcal{F}, \quad f(X) + f(Y) \le f(X + v) + f(Y - v).$$

With this v, we have $v \in \tilde{Y} \setminus \tilde{X}$ and

$$(\tilde{X} - u + v) \cap N = X + v \in \mathcal{F}, \qquad (\tilde{Y} + u - v) \cap N = Y - v \in \mathcal{F},$$

implying (65).

From valuated matroid to M$^\natural$-concave function. Let $g : \mathcal{B} \to \mathbb{R}$ be a valuated matroid defined on matroid bases $\mathcal{B}$. We define $\mathcal{F} \subseteq 2^N$ and a function $f : \mathcal{F} \to \mathbb{R}$ as follows:

$$\mathcal{F} = \{X \subseteq N \mid \exists Y \in \mathcal{B} \text{ s.t. } X \subseteq Y\}, \qquad f(X) = \max\{g(Y) \mid Y \supseteq X,\ Y \in \mathcal{B}\} \qquad (X \in \mathcal{F}).$$

Note that the restriction of f on $\mathcal{B}$ is equal to the original function g. Since $\mathcal{B}$ is the base family of a matroid, $\mathcal{F}$ is the independent set family of a matroid (see, e.g., [38, 41]). In the following, we show that f is an M$^\natural$-concave function.

For $X, Y \in \mathcal{F}$ and $u \in X \setminus Y$, we prove that either of (i) or (ii) in (M$^\natural$-EXC) holds. Let $\tilde{X}, \tilde{Y} \in \mathcal{B}$ be sets such that

$$X \subseteq \tilde{X}, \quad f(X) = g(\tilde{X}), \qquad Y \subseteq \tilde{Y}, \quad f(Y) = g(\tilde{Y}).$$

Note that $u \in X \subseteq \tilde{X}$.

[Case 1: $u \in \tilde{X} \setminus \tilde{Y}$] By the property (VM) of g, there exists some $v \in \tilde{Y} \setminus \tilde{X}$ such that

$$f(X) + f(Y) = g(\tilde{X}) + g(\tilde{Y}) \le g(\tilde{X} - u + v) + g(\tilde{Y} + u - v). \tag{66}$$

If $v \in Y$, then we have $v \in Y \setminus X$, $X - u + v \subseteq \tilde{X} - u + v$, and $Y + u - v \subseteq \tilde{Y} + u - v$, implying

$$f(X - u + v) \ge g(\tilde{X} - u + v), \qquad f(Y + u - v) \ge g(\tilde{Y} + u - v).$$

From this and (66) follows that the condition (ii) in (M$^\natural$-EXC) holds.

If $v \notin Y$, then we have $X - u \subseteq \tilde{X} - u + v$ and $Y + u \subseteq \tilde{Y} + u - v$, implying

$$f(X - u) \ge g(\tilde{X} - u + v), \qquad f(Y + u) \ge g(\tilde{Y} + u - v).$$

From this and (66) follows that the condition (i) in (M$^\natural$-EXC) holds.

[Case 2: $u \in \tilde{X} \cap \tilde{Y}$] We have $X - u \subseteq \tilde{X}$ and $Y + u \subseteq \tilde{Y}$, implying

$$f(X - u) \ge g(\tilde{X}) = f(X), \qquad f(Y + u) \ge g(\tilde{Y}) = f(Y).$$

Hence, the condition (i) in (M$^\natural$-EXC) holds.

This concludes the proof of M$^\natural$-concavity for the function f.

Appendix D. Algorithm for M$^\natural$-concave intersection problem. In this section, we consider the following problem called the *M$^\natural$-concave intersection problem*:

$$\textbf{(M}^\natural\textbf{I)} \qquad \text{Maximize } f_1(X) + f_2(X) \quad \text{subject to } X \in \mathcal{F}_1 \cap \mathcal{F}_2,$$

where $f_j : \mathcal{F}_j \to \mathbb{R}$ $(j = 1, 2)$ are M$^\natural$-concave functions defined on matroid independent sets $\mathcal{F}_j$. Recall that the Lagrangian relaxation problem (LR(λ)) in Section 4 is regarded as an M$^\natural$-concave intersection problem.

From the equivalence between M$^\natural$-concave functions and valuated matroids (see Section 2.4 and Appendix C), we see that the M$^\natural$-concave intersection problem can be reduced to the valuated matroid intersection problem formulated as follows:

$$\text{Maximize } g_1(X) + g_2(X) \quad \text{subject to } X \in \mathcal{B}_1 \cap \mathcal{B}_2,$$

where $g_j : \mathcal{B}_j \to \mathbb{R}$ $(j = 1, 2)$ are valuated matroids defined on matroid base families $\mathcal{B}_j$. The valuated matroid intersection problem is discussed in [30], and the results in the paper can be naturally restated in terms of the former problem. Indeed, we show in this section that the augmenting path algorithm proposed in [30] for the valuated matroid intersection problem is applicable to the problem (M$^\natural$I).

To solve the problem (M$^\natural$I), we consider a constrained problem (M$^\natural$I(k)) for each nonnegative integer k, which is the problem (M$^\natural$I) with an additional cardinality constraint $|X| = k$. It suffices to find an optimal solution to (M$^\natural$I(k)) for every k such that (M$^\natural$I(k)) has a feasible solution.

Optimal solutions to (M$^\natural$I(k)) can be found by an augmenting path algorithm with the aid of an auxiliary graph. Given $X \in \mathcal{F}_1 \cap \mathcal{F}_2$, we define an *auxiliary graph* $G_X = (V, A)$ associated with X by

$$\begin{aligned}
V &= N \cup \{s, t\},\\
A &= E_1 \cup E_2 \cup A_1 \cup A_2,\\
E_1 &= \{(u, v) \mid u \in X,\ v \in N \setminus X,\ X - u + v \in \mathcal{F}_1\},\\
E_2 &= \{(v, u) \mid v \in N \setminus X,\ u \in X,\ X + v - u \in \mathcal{F}_2\},\\
A_1 &= \{(s, v) \mid v \in N \setminus X,\ X + v \in \mathcal{F}_1\},\\
A_2 &= \{(v, t) \mid u \in N \setminus X,\ X + v \in \mathcal{F}_2\},
\end{aligned}$$

where s, t are new elements not in N. The arc length $\omega : A \to \mathbb{R}$ is defined by

$$\omega(a) = \begin{cases} f_1(X - u + v) - f_1(X) & (a = (u, v) \in E_1), \\ f_2(X + v - u) - f_2(X) & (a = (v, u) \in E_2), \\ f_1(X + v) - f_1(X) & (a = (s, v) \in A_1), \\ f_2(X + v) - f_2(X) & (a = (v, t) \in A_2). \end{cases}$$

Let $\bar{k}$ be the maximum integer such that $(\mathrm{M}^\natural\mathrm{I}(\bar{k}))$ has a feasible solution. Then, $(\mathrm{M}^\natural\mathrm{I}(k))$ has a feasible solution for each k with $0 \le k \le \bar{k}$ since $\mathcal{F}_1$ and $\mathcal{F}_2$ are matroid independent sets. Such $\bar{k}$ can be detected by using the following property.

Lemma D.1 (cf. [30, Lem. 3.1]) *Let $X \in \mathcal{F}_1 \cap \mathcal{F}_2$ be a feasible solution to* $(\mathrm{M}^\natural\mathrm{I}(k))$. *Then,* $(\mathrm{M}^\natural\mathrm{I}(k+1))$ *has a feasible solution if and only if there exists a directed path in G_X from s to t.*

It is easy to see that $X = \emptyset$ is an optimal solution to $(\mathrm{M}^\natural\mathrm{I}(0))$ since it is a unique feasible solution. The following property states that an optimal solution to $(\mathrm{M}^\natural\mathrm{I}(k+1))$ can be obtained by modification of an optimal solution to $(\mathrm{M}^\natural\mathrm{I}(k))$.

Lemma D.2 (cf. [30, Lem. 3.2]) *Let $X \in \mathcal{F}_1 \cap \mathcal{F}_2$ be an optimal solution to* $(\mathrm{M}^\natural\mathrm{I}(k))$*, and P be a longest directed path from s to t in G_X with respect to ω having the smallest number of arcs. Then, the set $\overline{X}$ defined by*

$$\overline{X} = X \setminus \{u \mid (u, v) \in P \cap E_1\} \cup \{v \mid (u, v) \in P \cap (E_1 \cup A_1)\}$$

is an optimal solution to $(\mathrm{M}^\natural\mathrm{I}(k+1))$.

The following lemma implies the existence of a longest directed path in the statement of Lemma D.2.

Lemma D.3 (cf. [30, Th. 5.2]) *A set $X \in \mathcal{F}_1 \cap \mathcal{F}_2$ with $|X| = k$ is an optimal solution to* $(\mathrm{M}^\natural\mathrm{I}(k))$ *if and only if G_X does not contain a directed cycle with positive length with respect to ω.*

Based on the lemmas above, we obtain the following augmenting path algorithm.

Augmenting Path Algorithm
Step 0: Set $X_0 := \emptyset$, $k := 0$.
Step 1: Construct the auxiliary graph G_{X_k}.
Step 2: If there exists no directed path in G_{X_k} from s to t, then stop.
Step 3: Find a longest path P from s to t in G_{X_k} having the smallest number of arcs.
Step 4: Output the set X_{k+1} given by

$$X_{k+1} := X_k \setminus \{u \mid (u, v) \in P \cap E_1\} \cup \{v \mid (u, v) \in P \cap (E_1 \cup A_1)\},$$

update k by $k := k + 1$, and go to Step 1.

By Lemma D.2, the set X_k is an optimal solution to $(\mathrm{M}^\natural\mathrm{I}(\bar{k}))$ for each k, and by Lemma D.3, the graph G_{X_k} does not contain a positive-length directed cycle. Hence, Step 3 can be done by using a shortest path algorithm. Hence, the algorithm can be implemented so that it runs in polynomial in n.

Theorem D.1 *The augmenting path algorithm finds optimal solutions X_k to the problems* $(\mathrm{M}^\natural\mathrm{I}(\bar{k}))$ *for all k with $0 \le k \le \bar{k}$ in time polynomial in n.*

We finally note that the augmenting algorithm can be implemented so that it applies comparison and addition operations to input numbers (i.e., no multiplication and division operations are used). This property is important in the computation of an optimal Lagrangian multiplier discussed in Appendix E.

Appendix E. Computing an optimal Lagrangian multiplier. In this section, we show that an optimal Lagrangian multiplier of (1BM$^\natural$I) can be computed in strongly-polynomial time. This can be shown by using Megiddo's parametric search technique as in [27, Sec. 2] (see also [39, Sec. 4.1]). Below we present the outline of the algorithm.

Recall that the Lagrangian relaxation (LR(λ)) of (1BM$^\natural$I) can be solved in strongly-polynomial time by using only comparison and addition operations (see Appendix D); we denote this algorithm as Algorithm A. To compute an optimal Lagrangian multiplier, we use a modified version of Algorithm A, denoted as Algorithm B. More precisely, Algorithm B computes an interval $[\ell, u]$ containing an optimal Lagrangian multiplier and a set $Z \in \mathcal{F}_1 \cap \mathcal{F}_2$ such that

$$z_{\mathrm{LR}}(\lambda) = f_1(Z) + f_2(Z) + \lambda(B - c(Z)) \ (\forall \lambda \in [\ell, u]). \tag{67}$$

The equation (67) implies that the function z_{LR} is linear in the interval $[\ell, u]$. Hence, we can compute an optimal Lagrangian multiplier λ_* easily since λ_* is a minimizer of function z_{LR}; indeed, at least one of ℓ and u is an optimal Lagrangian multiplier.

We initially set $\ell = -\infty$ and $u = +\infty$ in Algorithm B. In Algorithm B, we simulate the behavior of Algorithm A applied to the problem (LR(λ_*)), although we do not know the exact value of an optimal Lagrangian multiplier λ_* in advance. This means that λ_* is regarded as an unknown parameter, and addition and comparison operations are applied to linear functions with parameter λ_* in Algorithm B instead of real numbers as in Algorithm A. For example, if we add two linear functions $p\lambda_* + q$ and $r\lambda_* + s$, then we obtain a linear function $(p + r)\lambda_* + (q + s)$.

We then explain how to implement the comparison operation for two linear functions $p\lambda_* + q$ and $r\lambda_* + s$. As shown below, we can correctly determine if $p\lambda_* + q < r\lambda_* + s$, $p\lambda_* + q = r\lambda_* + s$, or $p\lambda_* + q > r\lambda_* + s$ holds, although we do not know the exact value of λ_*. As a byproduct of the comparison operation, we can also reduce the interval $[\ell, u]$ containing λ_* in some case.

If the two linear functions are the same, i.e., $p\lambda + q = r\lambda + s$ for all λ, then we have $p\lambda_* + q = r\lambda_* + s$. Hence, we assume that the two linear functions are distinct. If $p\lambda + q < r\lambda + s$ (resp., $p\lambda + q > r\lambda + s$) holds for all $\lambda \in [\ell, u]$, then we have $p\lambda_* + q < r\lambda_* + s$ (resp., $p\lambda_* + q > r\lambda_* + s$) since $\lambda_* \in [\ell, u]$. In either case, the interval $[\ell, u]$ remains the same.

We then consider the case where there exists a unique real number $\hat{\lambda} \in [\ell, u]$ such that $p\hat{\lambda} + q = r\hat{\lambda} + s$. In this case, comparison of two linear function reduces to comparison of $\hat{\lambda}$ and λ_*, i.e., we only need to check if $\hat{\lambda} < \lambda_*$, $\hat{\lambda} = \lambda_*$, or $\hat{\lambda} > \lambda_*$ holds. Since λ_* is a minimizer of the piecewise-linear convex function z_{LR}, we can easily determine the relation between $\hat{\lambda}$ and λ_* by using the left and right derivatives at $\hat{\lambda}$. Recall that the left derivative $(z_{\mathrm{LR}})'_+$ and the right derivative $(z_{\mathrm{LR}})'_-$ of the convex function z_{LR} at $\hat{\lambda}$ can be computed by solving (LR($\hat{\lambda} - \delta$)) and (LR($\hat{\lambda} + \delta$)) for a sufficiently small positive δ (see Lemma 4.1). The problems (LR($\hat{\lambda} + \delta$)) and (LR($\hat{\lambda} - \delta$)) can be solved in strongly-polynomial time by using Algorithm A twice. In this way, we can determine the relation between two linear functions in strongly-polynomial time. In addition, if $\hat{\lambda} = \lambda_*$ holds, then we stop Algorithm B by outputting $\hat{\lambda}$. Otherwise, we reduce the interval $[\ell, u]$ containing λ_* as follows:

$$\text{if } \hat{\lambda} < \lambda_*, \text{ then set } \ell := \max\{\ell, \hat{\lambda}\}, \qquad \text{if } \hat{\lambda} > \lambda_*, \text{ then set } u := \min\{u, \hat{\lambda}\}.$$

Suppose that Algorithm B terminates by outputting a set $Z \in \mathcal{F}_1 \cap \mathcal{F}_2$, as in Algorithm A. Since the addition and comparison operations are performed correctly, as explained above, we see that Z is an optimal solution to (LR(λ_*)). Moreover, we also see that Z is also an optimal solution to (LR(λ)) for all $\lambda \in [\ell, u]$, where $[\ell, u]$ is the interval at the end of the algorithm; this follows from the observation that for every $\lambda \in [\ell, u]$ the behavior of Algorithm A is the same as in the case with $\lambda = \lambda_*$. Hence, we obtain an interval $[\ell, u]$ and a set $Z \in \mathcal{F}_1 \cap \mathcal{F}_2$ satisfying desired conditions.