Budgeted Matroid Maximization: a Parameterized Viewpoint

Ilan Doron-Arad
Computer Science Department, Technion, Haifa, Israel

Ariel Kulik
CISPA Helmholtz Center for Information Security, Saarbrücken, Germany

Hadas Shachnai
Computer Science Department, Technion, Haifa, Israel

Abstract

We study budgeted variants of well known maximization problems with multiple matroid constraints. Given an ℓ-matchoid $\mathcal{M}$ on a ground set E, a profit function $p : E \to \mathbb{R}_{\geq 0}$, a cost function $c : E \to \mathbb{R}_{\geq 0}$, and a budget $B \in \mathbb{R}_{\geq 0}$, the goal is to find in the ℓ-matchoid a feasible set S of maximum profit $p(S)$ subject to the budget constraint, i.e., $c(S) \leq B$. The *budgeted ℓ-matchoid* (BM) problem includes as special cases budgeted ℓ-dimensional matching and budgeted ℓ-matroid intersection. A strong motivation for studying BM from parameterized viewpoint comes from the APX-hardness of unbudgeted ℓ-dimensional matching (i.e., $B = \infty$) already for $\ell = 3$. Nevertheless, while there are known FPT algorithms for the unbudgeted variants of the above problems, the *budgeted* variants are studied here for the first time through the lens of parameterized complexity.

We show that BM parametrized by solution size is $W[1]$-hard, already with a degenerate single matroid constraint. Thus, an exact parameterized algorithm is unlikely to exist, motivating the study of *FPT-approximation schemes* (FPAS). Our main result is an FPAS for BM (implying an FPAS for ℓ-dimensional matching and budgeted ℓ-matroid intersection), relying on the notion of representative set − a small cardinality subset of elements which preserves the optimum up to a small factor. We also give a lower bound on the minimum possible size of a representative set which can be computed in polynomial time.

2012 ACM Subject Classification Theory of computation

Keywords and phrases budgeted matching, budgeted matroid intersection, knapsack problems, FPT-approximation scheme

Digital Object Identifier 10.4230/LIPIcs.IPEC.2023.13

Related Version *Full Version*: https://arxiv.org/abs/2307.04173 [7]

Funding *Ariel Kulik*: Research supported by the European Reseach Concil (ERC) consolidator grant no. 725978 SYSTEMATICGRAPH.

1 Introduction

Numerous combinatorial optimization problems can be interpreted as *constrained budgeted problems*. In this setting, we are given a ground set E of elements and a family $\mathcal{I} \subseteq 2^E$ of subsets of E known as the *feasible sets*. We are also given a cost function $c : E \to \mathbb{R}$, a profit function $p : E \to \mathbb{R}$, and a budget $B \in \mathbb{R}$. A *solution* is a feasible set $S \in \mathcal{I}$ of bounded cost $c(S) \leq B$.[1] Broadly speaking, the goal is to find a solution S of maximum profit. Notable examples include budgeted matching [1] and budgeted matroid intersection [3, 20], shortest weight-constrained path [18], and constrained minimum spanning trees [36].

[1] For a function $f : A \to \mathbb{R}$ and a subset of elements $C \subseteq A$, define $f(C) = \sum_{e \in C} f(e)$.

$$\mathcal{I}_1 = \{\{\emptyset\}, \{\mathbf{a}\}, \{\mathbf{b}\}, \{\mathbf{c}\}, \{\mathbf{a}, \mathbf{b}\}, \{\mathbf{c}, \mathbf{b}\}\}$$
$$\mathcal{I}_2 = \{\{\emptyset\}, \{\mathbf{a}\}, \{\mathbf{b}\}, \{\mathbf{c}\}, \{\mathbf{a}, \mathbf{b}\}, \{\mathbf{a}, \mathbf{c}\}\}$$
$$\mathcal{I}_3 = \{\{\emptyset\}, \{\mathbf{a}\}, \{\mathbf{b}\}, \{\mathbf{c}\}, \{\mathbf{a}, \mathbf{b}\}, \{\mathbf{b}, \mathbf{c}\}\}$$
$$\mathcal{I} = \mathcal{I}_1 \cap \mathcal{I}_2 \cap \mathcal{I}_3 = \{\{\emptyset\}, \{\mathbf{a}\}, \{\mathbf{b}\}, \{\mathbf{c}\}, \{\mathbf{a}, \mathbf{b}\}\}$$

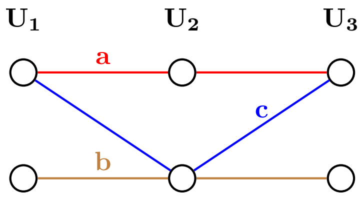

Figure 1 A 3-dimensional matching viewed as a 3-matroid intersection. Each element in $E \subset U_1 \times U_2 \times U_3$ is represented by a path of three vertices, where the paths are distinguished by different colors, that is $E = \{a, b, c\}$. There is a matroid constraint $(E, \mathcal{I}_i)$ for each $U_i, i = 1, 2, 3$. The feasible sets $\mathcal{I}$ for the matching are exactly the common independent sets of $\mathcal{I}_i, i = 1, 2, 3$.

Despite the wide interest in constrained budgeted problems in approximation algorithms, not much is known about this intriguing family of problems in terms of parameterized complexity. In this work, we study budgeted maximization with the fairly general *ℓ-dimensional matching*, *ℓ-matroid intersection*, and *ℓ-matchoid* constraints.

An ℓ-dimensional matching constraint is a set system $(E, \mathcal{I})$, where $E \subseteq U_1 \times \ldots \times U_\ell$ for ℓ sets $U_1, \ldots, U_\ell$. The feasible sets $\mathcal{I}$ are all subsets $S \subseteq E$ which satisfy the following. For any two distinct tuples $(e_1, \ldots, e_\ell), (f_1, \ldots, f_\ell) \in S$ and every $i \in [\ell]$ it holds that $e_i \neq f_i$.[2] Informally, the input for *budgeted ℓ-dimensional matching* is an ℓ-dimensional matching constraint $(E, \mathcal{I})$, profits and costs for the elements in E, and a budget. The objective is to find a feasible set which maximizes the profit subject to the budget constraint (see below the formal definition).

We now define an ℓ-matroid intersection. A *matroid* is a set system $(E, \mathcal{I})$, where E is a finite set and $\mathcal{I} \subseteq 2^E$, such that

- $\emptyset \in \mathcal{I}$.
- The *hereditary property*: for all $A \in \mathcal{I}$ and $B \subseteq A$ it holds that $B \in \mathcal{I}$.
- The *exchange property*: for all $A, B \in \mathcal{I}$ where $|A| > |B|$ there is $e \in A \setminus B$ such that $B \cup \{e\} \in \mathcal{I}$.

For a fixed $\ell \geq 1$, let $(E, \mathcal{I}_1), (E, \mathcal{I}_2), \ldots, (E, \mathcal{I}_\ell)$ be ℓ matroids on the same ground set E. An ℓ-matroid intersection is a set system $(E, \mathcal{I})$ where $\mathcal{I} = \mathcal{I}_1 \cap \mathcal{I}_2 \cap \ldots \cap \mathcal{I}_\ell$. Observe that ℓ-dimensional matching, where $E \subseteq U_1 \times \ldots \times U_\ell$, is a special case of ℓ-matroid intersection: For each $i \in [\ell]$, define a partition matroid $(E, \mathcal{I}_i)$, where any feasible set $S \in \mathcal{I}_i$ may contain each element $e \in U_i$ in the i-th coordinate at most once, i.e.,

$$\mathcal{I}_i = \{S \subseteq E \mid \forall (e_1, \ldots, e_\ell) \neq (f_1, \ldots, f_\ell) \in S : e_i \neq f_i\}.$$

We give an illustration in Figure 1. It can be shown that $(E, \mathcal{I}_i)$ is a matroid for all $i \in \ell$ (see, e.g., [37]).

The above constraint families can be generalized to the notion of *ℓ-matchoid*. Informally, an ℓ-matchoid is an intersection of an unbounded number of matroids, where each element belongs to at most ℓ of the matroids. Formally, for any $\ell \geq 1$, an ℓ-matchoid on a set E is a collection $\mathcal{M} = \{M_i = (E_i, \mathcal{I}_i)\}_{i \in [s]}$ of $s \in \mathbb{N}$ matroids, where for each $i \in [s]$ it holds that $E_i \subseteq E$, and every $e \in E$ belongs to at most ℓ sets in $\{E_1, \ldots, E_s\}$, i.e., $|\{i \in [s] \mid e \in E_i\}| \leq \ell$. A set $S \subseteq E$ is *feasible* for $\mathcal{M}$ if for all $i \in [s]$ it holds that $S \cap E_i \in \mathcal{I}_i$. Let $\mathcal{I}(\mathcal{M}) = \{S \subseteq E \mid \forall i \in [s] : S \cap E_i \in \mathcal{I}_i\}$ be all feasible sets of $\mathcal{M}$. For all $k \in \mathbb{N}$, we

[2] For any $k \in \mathbb{N}$ let $[k] = \{1, 2, \ldots, k\}$.

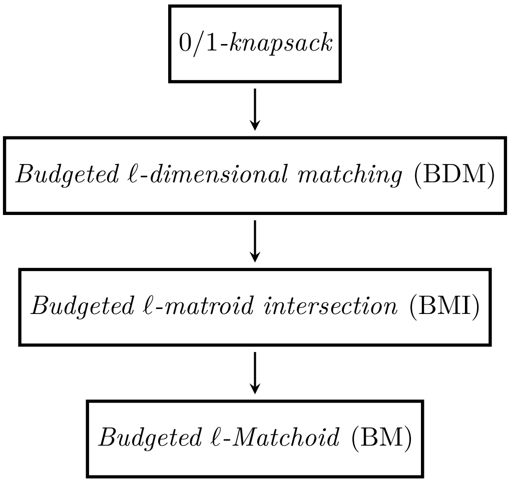

Figure 2 An overview of constrained budgeted problems. An arrow from problem A to problem B indicates that A is a special case of B.

use $\mathcal{M}_k \subseteq \mathcal{I}(\mathcal{M})$ to denote all feasible sets of $\mathcal{M}$ of cardinality at most k. Clearly, ℓ-matroid intersection (and also ℓ-dimensional matching) is the special case of ℓ-matchoid where the $s(=\ell)$ matroids are defined over the same ground set E.

In the *budgeted ℓ-matchoid (BM)* problem, we are given an ℓ-matchoid along with a cost function, profit function, and a budget; our goal is to maximize the profit of a feasible set under the budget constraint. The *budgeted ℓ-matroid intersection (BMI)* and *budgeted ℓ-dimensional matching (BDM)* are the special cases where the ℓ-matchoid is an ℓ-matroid intersection and ℓ-dimensional matching, respectively. Each of these problems generalizes the classic 0/1-knapsack, where all sets are feasible. Figure 2 shows the relations between the problems. Henceforth, we focus on the BM problem.

Formally, a BM instance is a tuple $I = (E, \mathcal{M}, c, p, B, k, \ell)$, where E is a ground set of elements, $\mathcal{M}$ is an ℓ-matchoid on E, $c : E \to \mathbb{N}_{>0}$ is a cost function, $p : E \to \mathbb{N}_{>0}$ is a profit function, $B \in \mathbb{N}_{>0}$ is a budget, and $k, \ell \in \mathbb{N}_{>0}$ are integer parameters.[3] In addition, each matroid $(E_i, \mathcal{I}_i) \in \mathcal{M}$ has a *membership oracle*, which tests whether a given subset of E_i belongs to $\mathcal{I}_i$ or not in a single query. Indeed, with no membership oracle, the representation size for a matroid over n elements may be exponential in n. A *solution* of I is a feasible set $S \in \mathcal{M}_k$ such that $c(S) \leq B$. The objective is to find a solution S of I such that $p(S)$ is maximized. We consider algorithms parameterized by k and ℓ (equivalently, $k + \ell$).

We note that even with no budget constraint (i.e., $c(E) \leq B$), where the ℓ-matchoid is restricted to be a 3-dimensional matching, BM is MAX SNP-complete [26], i.e., it cannot admit a *polynomial time approximation scheme (PTAS)* unless P=NP. On the other hand, the ℓ-dimensional matching and even the ℓ-matchoid problem (without a budget), parameterized by ℓ and the solution size k, are *fixed parameter tractable (FPT)* [19, 22]. This motivates our study of BM through the lens of parameterized complexity. We first observe that BM parameterized by the solution size is W[1]-hard, already with a *uniform* matroid where all sets are feasible (i.e., knapsack parametrized by the cardinality of the solution, k).

▶ **Theorem 1.** BM *is* $W[1]$-hard*.*

[3] We assume integral values for simplicity; our results can be generalized also for real values.

By the hardness result in Theorem 1, the best we can expect for BM in terms of parametrized algorithms, is an *FPT-approximation scheme (FPAS)*. An FPAS with parameterization κ for a maximization problem Π is an algorithm whose input is an instance I of Π and an $\varepsilon > 0$, which produces a solution S of I of value $(1-\varepsilon) \cdot \mathrm{OPT}(I)$ in time $f(\varepsilon, \kappa(|I|)) \cdot |I|^{O(1)}$ for some computable function f, where $|I|$ denotes the encoding size of I and $\mathrm{OPT}(I)$ is the optimum value of I. We refer the reader to [34, 14] for comprehensive surveys on parameterized approximation schemes and parameterized approximations in general. To derive an FPAS for BM, we use a small cardinality *representative set*, which is a subset of elements containing the elements of an *almost* optimal solution for the instance. The representative set has a cardinality depending solely on $\ell, k, \varepsilon^{-1}$ and is constructed in FPT time. Formally,

▶ **Definition 2.** *Let $I = (E, \mathcal{M}, c, p, B, k, \ell)$ be a* BM *instance, $0 < \varepsilon < \frac{1}{2}$ and $R \subseteq E$. Then R is a* representative set *of I and ε if there is a solution S of I such that the following holds.*
1. $S \subseteq R$.
2. $p(S) \geq (1 - 2\varepsilon) \cdot \mathrm{OPT}(I)$.

We remark that Definition 2 slightly resembles the definition of *lossy kernel* [31]. Nonetheless, the definition of lossy kernel does not apply to problems in the oracle model, including BM (see Section 6 for further details).

The main technical contribution of this paper is the design of a small cardinality representative set for BM. Our representative set is constructed by forming a collection of $f(\ell, k, \varepsilon^{-1})$ *profit classes*, where the elements of each profit class have roughly the same profit. Then, to construct a representative set for the instance, we define a residual problem for each profit class which enables to circumvent the budget constraint. These residual problems can be solved efficiently using a construction of [22]. We show that combining the solutions for the residual problems, we obtain a representative set. In the following, we use $\tilde{O}(n)$ for $O(n \cdot \mathrm{poly}(\log(n)))$.

▶ **Theorem 3.** *There is an algorithm that given a* BM *instance $I = (E, \mathcal{M}, c, p, B, k, \ell)$ and $0 < \varepsilon < \frac{1}{2}$, returns in time $|I|^{O(1)}$ a representative set $R \subseteq E$ of I and ε such that $|R| = \tilde{O}\left(\ell^{(k-1)\cdot \ell} \cdot k^2 \cdot \varepsilon^{-2}\right)$.*

Given a small cardinality representative set, it is easy to derive an FPAS. Specifically, using an exhaustive enumeration over the representative set as stated in Theorem 3, we can construct the following FPAS for BM, which naturally applies also for BMI and BDM.

▶ **Theorem 4.** *For any* BM *instance $I = (E, \mathcal{M}, c, p, B, k, \ell)$ and $0 < \varepsilon < \frac{1}{2}$, there is an* FPAS *whose running time is $|I|^{O(1)} \cdot \tilde{O}\left(\ell^{k^2 \cdot \ell} \cdot k^{O(k)} \cdot \varepsilon^{-2k}\right)$.*

Our FPAS cannot be significantly improved even for very restricted ℓ-matchoids. Namely, even if the ℓ-matchoid is a single matroid there cannot be an FPAS for BMI with running time polynomial in $\frac{1}{\varepsilon}$, where ε is the given error parameter [11]. To complement the above construction of a representative set, we show that even for the special case of an ℓ-dimensional matching constraint, it is unlikely that a representative set of significantly smaller cardinality can be constructed in polynomial time. The next result applies to the special case of BDM.

▶ **Theorem 5.** *For any function $f : \mathbb{N} \to \mathbb{N}$, and $c_1, c_2 \in \mathbb{R}$ such that $c_2 - c_1 < 0$, there is no algorithm which finds for a given* BM *instance $I = (E, \mathcal{M}, c, p, B, k, \ell)$ and $0 < \varepsilon < \frac{1}{2}$ a representative set of size $O\left(f(\ell) \cdot k^{\ell - c_1} \cdot \frac{1}{\varepsilon^{c_2}}\right)$ of I and ε in time $|I|^{O(1)}$, unless* coNP $\subseteq$ NP/poly.

In the proof of Theorem 1, we use a lower bound on the kernel size of the *Perfect* 3-*Dimensional Matching* (3-PDM) problem, due to Dell and Marx [5, 6].[4] In our hardness result, we are able to efficiently construct a kernel for 3-PDM using a representative set for BM, already for the special case of 3-dimensional matching constraint, uniform costs, and uniform profits.

Related Work

While BM is studied here for the first time, special cases of the problem have been extensively studied from both parameterized and approximative points of view. For maximum weighted ℓ-matchoid without a budget constraint, Huang and Ward [22] obtained a deterministic FPT algorithm, and algorithms for a more general problem, involving a *coverage function* objective rather than a linear objective. Their result differentiates the ℓ-matchoid problem from the matroid ℓ-parity problem which cannot have an FPT algorithm in general matroids [32, 24]. Interestingly, when the matroids are given a linear representation, the matroid ℓ-parity problem admits a randomized FPT algorithm [35, 15] and a deterministic FPT algorithm [30]. We use a construction of [22] as a building block of our algorithm.

The ℓ-dimensional k-matching problem (i.e., the version of the problem with no budget parametrized by k and ℓ) has received considerable attention in previous studies. Goyal et al. [19] presented a deterministic FPT algorithm whose running time is $O^*(2.851^{(\ell-1)\cdot k})$ for the weighted version of ℓ-dimensional k-matching, where O^* is used to suppress polynomial factor in the running time. This result improves a previous result of [4]. For the unweighted version of ℓ-dimensional k-matching, the state of the art is a randomized FPT algorithm with running time $O^*(2^{(\ell-2)\cdot k})$ [2], improving a previous result for the problem [27].

Budgeted problems are well studied in approximation algorithms. As BM is a generalization of classic 0/1-knapsack, it is known to be NP-hard. However, while knapsack admits a *fully PTAS (FPTAS)* [33], BM is unlikely to admit a PTAS, even for the special case of 3-dimensional matching with no budget constraint [26]. Consequently, there has been extensive research work to identify special cases of BM which admit approximation schemes.

For the budgeted matroid independent set (i.e., the special case of BM where the ℓ-matchoid consists of a single matroid), Doron-Arad et al. [9] developed an *efficient PTAS (EPTAS)* using the representative set based technique. This algorithm was later generalized in [8] to tackle budgeted matroid intersection and budgeted matching (both are special cases of BM where $\ell = 2$), improving upon a result of Berger et al. [1]. For the special case where the matroid is a *laminar matroid*, there is an FPTAS [10]. We generalize some of the technical ideas of [9, 8] to the setting of ℓ-matchoid and parametrized approximations.

Organization of the paper. Section 2 describes our construction of a representative set. In Section 3 we present our FPAS for BM. Section 4 contains the proofs of the hardness results given in Theorem 1 and in Theorem 1. In Section 5 we present an auxiliary approximation algorithm for BM. We conclude in Section 6 with a summary and some directions for future work.

[4] We refer the reader e.g., to [16], for the formal definition of kernels.

2 Representative Set

In this section we construct a representative set for BM. Our first step is to round the profits of a given instance, and to determine the low profit elements that can be discarded without incurring significant loss of profit. We find a small cardinality representative set from which an almost optimal solution can be selected via enumeration yielding an FPAS (see Section 3).

We proceed to construct a representative set whose cardinality depends only on ε^{-1}, k, and ℓ. This requires the definition of *profit classes*, namely, a partition of the elements into groups, where the elements in each group have similar profits. Constructing a representative set using this method requires an approximation of the optimum value of the input BM instance I. To this end, we use a $\frac{1}{2\ell}$-approximation $\alpha = \mathsf{ApproxBM}(I)$ of the optimum value $\mathrm{OPT}(I)$ described below.

▶ **Lemma 6.** *Given a* BM *instance* $I = (E, \mathcal{M}, c, p, B, k, \ell)$*, there is an algorithm* ApproxBM *which returns in time* $|I|^{O(1)}$ *a value* α *such that* $\frac{\mathrm{OPT}(I)}{2\ell} \leq \alpha \leq \mathrm{OPT}(I)$*.*

The proof of Lemma 6 is given in Section 5. The proof utilizes a known approximation algorithm for the unbudgeted version of BM [23, 25] which is then transformed into an approximation algorithm for BM using a technique of [29].

The first step in designing the profit classes is to determine a set of *profitable* elements. required for obtaining an almost optimal solution. This set allows us to

construct only a small number of profit classes. We define the set of *profitable* elements w.r.t. I, α, and ε as

$$H[I, \alpha, \varepsilon] = \left\{ e \in E \mid \frac{\varepsilon \cdot \alpha}{k} < p(e) \leq 2 \cdot \ell \cdot \alpha \right\}. \tag{1}$$

The upper bound on $p(e)$ is purely technical; it is used later to upper bound the number of profit classes. When clear from the context, we simply use $H = H[I, \alpha, \varepsilon]$. Consider the non-profitable elements. The next lemma states that omitting these elements indeed has small effect on the profit of the solution set.

▶ **Lemma 7.** *For every* BM *instance* $I = (E, \mathcal{C}, c, p, B, k, \ell)$*,* $\frac{\mathrm{OPT}(I)}{2\ell} \leq \alpha \leq \mathrm{OPT}(I)$*,* $0 < \varepsilon < \frac{1}{2}$*, and* $S \in \mathcal{M}_k$ *it holds that* $p\left(S \setminus H[I, \alpha, \varepsilon]\right) \leq \varepsilon \cdot \mathrm{OPT}(I)$*.*

Proof. We note that

$$p\left(S \setminus H[I, \alpha(I), \varepsilon]\right) \leq k \cdot \frac{\varepsilon \cdot \alpha}{k} = \varepsilon \cdot \alpha \leq \varepsilon \cdot \mathrm{OPT}(I).$$

The first inequality holds since each element in $S \setminus H[I, \alpha(I), \varepsilon]$ has profit at most $\frac{\varepsilon \cdot \alpha}{k}$ by (1); in addition, since $S \in \mathcal{M}_k$ it follows that S contains at most k elements. The second inequality holds as $\alpha \leq \mathrm{OPT}(I)$. ◀

Using Lemma 7, our representative set can be constructed exclusively from the profitable elements. We can now partition the profitable elements into a small number of *profit classes*. There is a profit class $\mathcal{K}_r(\alpha)$ for a suitable range of profit *boundaries* r and a constant factor approximation α for OPT. Specifically, let

$$D(I, \varepsilon) = \left\{ r \in \mathbb{N}_{>0} \mid (1-\varepsilon)^{r-1} \geq \frac{\varepsilon}{2 \cdot \ell \cdot k} \right\} \tag{2}$$

be the set of *boundaries* for the profit classes (defined below). We simplify by $D = D(I, \varepsilon)$. For all $r \in D$, and $\frac{\mathrm{OPT}(I)}{2\ell} \leq \alpha \leq \mathrm{OPT}(I)$, define the *$r$-profit class* as

$$\mathcal{K}_r(\alpha) = \left\{ e \in E \;\middle|\; \frac{p(e)}{2 \cdot \ell \cdot \alpha} \in \left((1-\varepsilon)^r, (1-\varepsilon)^{r-1}\right] \right\}. \tag{3}$$

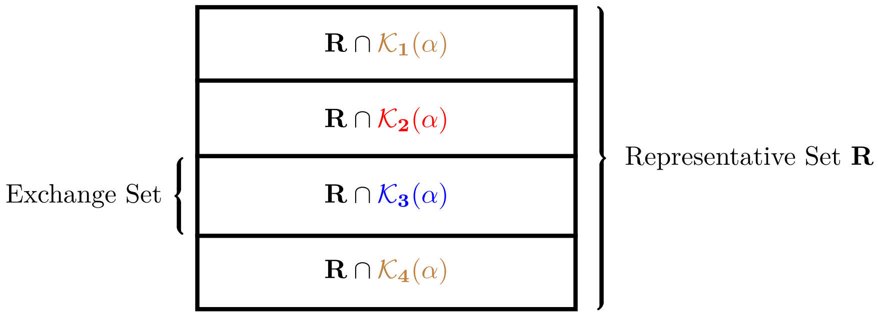

Figure 3 An illustration of our construction of a representative set R, using a union of exchange sets, one for each profit class $\mathcal{K}_1(\alpha)$, $\mathcal{K}_2(\alpha)$, $\mathcal{K}_3(\alpha)$, $\mathcal{K}_4(\alpha)$.

In words, each profit class $r \in D$ contains profitable elements (and may contain some elements that are *almost* profitable due to our $\frac{1}{2\ell}$-approximation for $\mathrm{OPT}(I)$), where the profits of any two elements that belong to the r-profit class can differ by at most a multiplicative factor of $(1-\varepsilon)$. We use the following simple upper bound on the number of profit classes.

▶ **Lemma 8.** *For every* BM *instance I and $0 < \varepsilon < \frac{1}{2}$ there are $O(k \cdot \ell \cdot \varepsilon^{-2})$ profit classes.*

Proof. We note that

$$\log_{1-\varepsilon}\left(\frac{\varepsilon}{2\ell \cdot k}\right) \leq \frac{\ln\left(\frac{2\ell \cdot k}{\varepsilon}\right)}{-\ln(1-\varepsilon)} \leq \frac{2\ell \cdot k \cdot \varepsilon^{-1}}{\varepsilon}. \tag{4}$$

The second inequality follows from $x < -\ln(1-x), \forall x > -1, x \neq 0$, and $\ln(y) < y, \forall y > 0$. By (2) the number of profit classes is bounded by

$$|D| \leq \log_{1-\varepsilon}\left(\frac{\varepsilon}{2\ell \cdot k}\right) + 1 = O(k \cdot \ell \cdot \varepsilon^{-2}). \tag{5}$$

The last inequality follows from (4). ◀

Next, we define an *exchange set* for each profit class. This facilitates the construction of a representative set. Intuitively, a subset of elements X forms an exchange set for a profit class $\mathcal{K}_r(\alpha)$ if any feasible set Δ and element $a \in (\Delta \cap \mathcal{K}_r(\alpha)) \setminus X$ can be replaced (while maintaining feasibility) by some element $b \in (X \cap \mathcal{K}_r(\alpha)) \setminus \Delta$ such that the cost of b is upper bounded by the cost of a. Formally,

▶ **Definition 9.** *Let $I = (E, \mathcal{M}, c, p, B, k, \ell)$ be a* BM *instance, $0 < \varepsilon < \frac{1}{2}$, $\frac{\mathrm{OPT}(I)}{2\ell} \leq \alpha \leq \mathrm{OPT}(I)$, $r \in D(I, \varepsilon)$, and $X \subseteq \mathcal{K}_r(\alpha)$. We say that X is an* exchange set *for I, ε, α, and r if: For all $\Delta \in \mathcal{M}_k$ and $a \in (\Delta \cap \mathcal{K}_r(\alpha)) \setminus X$ there is $b \in (\mathcal{K}_r(\alpha) \cap X) \setminus \Delta$ satisfying*
- *$c(b) \leq c(a)$.*
- *$\Delta - a + b \in \mathcal{M}_k$.*

The key argument in this section is that if a set $R \subseteq E$ satisfies that $R \cap \mathcal{K}_r(\alpha)$ is an exchange set for any $r \in D$, then R is a representative set. This allows us to construct a representative set using a union of disjoint exchange sets, one for each profit class. We give an illustration in Figure 3.

▶ **Lemma 10.** *Let $I = (E, \mathcal{M}, c, p, B, k, \ell)$ be a* BM *instance, $0 < \varepsilon < \frac{1}{2}$, $\frac{\mathrm{OPT}(I)}{2\ell} \leq \alpha \leq \mathrm{OPT}(I)$, and $R \subseteq E$. If for all $r \in D = D(I, \varepsilon)$ it holds that $R \cap \mathcal{K}_r(\alpha)$ is an exchange set for I, ε, α, and r, then R is a representative set of I and ε.*

For the proof of Lemma 10, we define a *substitution* of some feasible set $G \in \mathcal{M}_k$. We will use G later only as an optimal solution; however, we can state the following claims for a general $G \in \mathcal{M}_k$. We require that a substitution preserves the number of profitable elements in G from each profit class, so a substitution guarantees a profit similar to the profit of G.

▶ **Definition 11.** *For $G \in \mathcal{M}_k$ and $Z_G \subseteq \bigcup_{r \in D} \mathcal{K}_r(\alpha)$, we say that Z_G is a* substitution *of G if the following holds.*

1. $Z_G \in \mathcal{M}_k$.
2. $c(Z_G) \leq c(G)$.
3. *For all $r \in D$ it holds that $|\mathcal{K}_r(\alpha) \cap Z_G| = |\mathcal{K}_r(\alpha) \cap G|$.*

Proof of Lemma 10. We first show that every set $G \in \mathcal{M}_k$ has a substitution which is a subset of R.

▷ Claim 12. For any $G \in \mathcal{M}_k$ there is a substitution Z_G of G such that $Z_G \subseteq R$.

Proof. Let $G \in \mathcal{M}_k$ and let Z_G be a substitution of G such that $|Z_G \cap R|$ is maximal among all substitutions of G; formally, let $\mathcal{S}(G)$ be all substitutions of G and let

$$Z_G \in \{Z \in \mathcal{S}(G) \mid |Z \cap R| = \max_{Z' \in \mathcal{S}(G)} |Z' \cap R|\}.$$

Since $G \cap \bigcup_{r \in D} \mathcal{K}_r(\alpha)$ is in particular a substitution of G it follows that $\mathcal{S}(G) \neq \emptyset$; thus, Z_G is well defined. Assume towards a contradiction that there is $a \in Z_G \setminus R$; then, by Definition 11 there is $r \in D$ such that $a \in \mathcal{K}_r(\alpha)$. Because $R \cap \mathcal{K}_r(\alpha)$ is an exchange set for I, ε, α, and r, by Definition 9 there is $b \in (\mathcal{K}_r(\alpha) \cap R) \setminus Z_G$ such that $c(b) \leq c(a)$ and $Z_G - a + b \in \mathcal{M}_k$. Then, the properties of Definition 11 are satisfied for $Z_G - a + b$ by the following.

1. $Z_G - a + b \in \mathcal{M}_k$ by the definition of b.
2. $c(Z_G - a + b) \leq c(Z_G) \leq c(G)$ because $c(b) \leq c(a)$.
3. for all $r' \in D$ it holds that $|\mathcal{K}_{r'}(\alpha) \cap (Z_G - a + b)| = |\mathcal{K}_{r'}(\alpha) \cap Z_G| = |\mathcal{K}_{r'}(\alpha) \cap G|$ because $a, b \in \mathcal{K}_r(\alpha)$.

By the above, and using and Definition 11, we have that $Z_G + a - b$ is a substitution of G; that is, $Z_G + a - b \in \mathcal{S}(G)$. Moreover,

$$|R \cap (Z_G - a + b)| > |R \cap Z_G| = \max_{Z \in \mathcal{S}(G)} |Z \cap R|. \tag{6}$$

The first inequality holds since $a \in Z_G \setminus R$ and $b \in R$. Thus, we have found a substitution of G which contains more elements in R than $Z_G \in \mathcal{S}(G)$. A contradiction to the definition of Z_G as a substitution of G having a maximum number of elements in R. Hence, $Z_G \subseteq R$, as required. ◁

Let G be an optimal solution for I. We complete the proof of Lemma 10 by showing that a substitution of G which is a subset of R yields a profit at least $(1 - 2\varepsilon) \cdot \mathrm{OPT}(I)$. Let $H[I, \alpha, \varepsilon] = H$ be the set of profitable elements w.r.t. I, α and ε (as defined in (1)). By Claim 12, as $G \in \mathcal{M}_k$, it has a substitution $Z_G \subseteq R$. Then,

$$\begin{aligned} p(Z_G) &\geq \sum_{r \in D} p(\mathcal{K}_r(\alpha) \cap Z_G) \\ &\geq \sum_{r \in D \text{ s.t. } \mathcal{K}_r(\alpha) \neq \emptyset} |\mathcal{K}_r(\alpha) \cap Z_G| \cdot \min_{e \in \mathcal{K}_r(\alpha)} p(e) \\ &\geq \sum_{r \in D \text{ s.t. } \mathcal{K}_r(\alpha) \neq \emptyset} |\mathcal{K}_r(\alpha) \cap G| \cdot (1 - \varepsilon) \cdot \max_{e \in \mathcal{K}_r(\alpha)} p(e) \\ &\geq (1 - \varepsilon) \cdot p(G \cap H). \end{aligned} \tag{7}$$

The third inequality follows from (3), and from Property 3 in Definition 11. The last inequality holds since for every $e \in H$ there is $r \in D$ such that $e \in \mathcal{K}_r(\alpha)$, by (1) and (3). Therefore,

$$\begin{aligned} p(Z_G) &\geq (1-\varepsilon) \cdot p(G \cap H) \\ &= (1-\varepsilon) \cdot (p(G) - p(G \setminus H)) \\ &\geq (1-\varepsilon) \cdot p(G) - p(G \setminus H) \\ &\geq (1-\varepsilon) \cdot p(G) - \varepsilon \cdot \mathrm{OPT}(I) \\ &= (1-\varepsilon) \cdot \mathrm{OPT}(I) - \varepsilon \cdot \mathrm{OPT}(I) \\ &= (1-2\varepsilon) \cdot \mathrm{OPT}(I). \end{aligned} \tag{8}$$

The first inequality follows from (7). The last inequality holds by Lemma 7. The second equality holds since G is an optimal solution for I. To conclude, by Properties 1 and 2 in Definition 11, it holds that $Z_G \in \mathcal{M}_k$, and $c(Z_G) \leq c(G) \leq B$; thus, Z_G is a a solution for I. Also, by (8), it holds that $p(Z_G) \geq (1-2\varepsilon) \cdot \mathrm{OPT}(I)$ as required (see Definition 2). ◀

By Lemma 10, our end goal of constructing a representative set is reduced to efficiently finding exchange sets for all profit classes. This can be achieved by the following result, which is a direct consequence of Theorem 3.6 in [22]. As the result of [22] refers to a *maximization* version of exchange sets, we first present an analogue to Definition 9 for maximization exchange sets (as in [22]), using our notation.

▶ **Definition 13.** *Let* $I = (E, \mathcal{M}, c, p, B, k, \ell)$ *be a* BM *instance,* $0 < \varepsilon < \frac{1}{2}$, $\frac{\mathrm{OPT}(I)}{2\ell} \leq \alpha \leq \mathrm{OPT}(I)$, $r \in D(I, \varepsilon)$, *and* $X \subseteq \mathcal{K}_r(\alpha)$. *We say that* X *is a* maximization exchange set *for* I, ε, α, *and* r *if: For all* $\Delta \in \mathcal{M}_k$ *and* $a \in (\Delta \cap \mathcal{K}_r(\alpha)) \setminus X$ *there is* $b \in (\mathcal{K}_r(\alpha) \cap X) \setminus \Delta$ *satisfying*

- $c(a) \leq c(b)$.
- $\Delta - a + b \in \mathcal{M}_k$.

We remark that the same construction and the proof of Theorem 3.6 in [21] hold for our exchange sets (in Definition 9) as well. Hence, we have the following.

▶ **Lemma 14.** *Given a* BM *instance* $I = (E, \mathcal{M}, c, p, B, k, \ell)$, $0 < \varepsilon < \frac{1}{2}$, $\frac{\mathrm{OPT}(I)}{2\ell} \leq \alpha \leq \mathrm{OPT}(I)$, *and* $r \in D(I, \varepsilon)$, *there is an algorithm* ExSet *which returns in time* $\tilde{O}\left(\ell^{(k-1)\cdot\ell} \cdot k\right) \cdot |I|^{O(1)}$ *an exchange set* X *for* I, ε, α, *and* r, *such that* $|X| = \tilde{O}\left(\ell^{(k-1)\cdot\ell} \cdot k\right)$.

Algorithm 1 RepSet($I = (E, \mathcal{M}, c, p, B, k, \ell), \varepsilon$).

input : A BM instance I, and an error parameter $0 < \varepsilon < \frac{1}{2}$.
output : A representative set R of I and ε.
if $\ell^{(k-1)\cdot\ell} \cdot k^2 \cdot \varepsilon^{-2} > |I|$ **then**
  Return E
Compute $\alpha \leftarrow$ ApproxBM(I).
for $r \in D(I, \varepsilon)$ **do**
  $R \leftarrow R \cup$ ExSet($I, \varepsilon, \alpha, r$).
Return R

Using Lemmas 10 and 14, a representative set of I can be constructed as follows. If the parameters ℓ and k are too high w.r.t. $|I|$, return the trivial representative set E in polynomial time. Otherwise, compute an approximation for $\mathrm{OPT}(I)$, and define the profit classes. Then, the representative set is constructed by finding an exchange set for each profit class. The pseudocode of the algorithm is given in Algorithm 1.

▶ **Lemma 15.** *Given a* BM *instance* $I = (E, \mathcal{M}, c, p, B, k, \ell)$*, and* $0 < \varepsilon < \frac{1}{2}$*, Algorithm 1 returns in time* $|I|^{O(1)}$ *a representative set* $R \subseteq E$ *of* I *and* ε *such that* $|R| = \tilde{O}\left(\ell^{(k-1)\cdot\ell} \cdot k^2 \cdot \varepsilon^{-2}\right)$.

Proof. Clearly, if $\ell^{(k-1)\cdot\ell} \cdot k^2 \cdot \varepsilon^{-2} > |I|$, then by Step 2 the algorithm runs in time $|I|^{O(1)}$ and returns the trivial representative set E. Thus, we may assume below that $\ell^{(k-1)\cdot\ell} \cdot k^2 \cdot \varepsilon^{-2} \leq |I|$. The running time of Step 3 is $|I|^{O(1)}$ by Lemma 6. Each iteration of the **for** loop in Step 4 can be computed in time $\tilde{O}(\ell^{(k-1)\cdot\ell} \cdot k) \cdot |I|^{O(1)}$, by Lemma 14. Hence, as we have $|D| = |D(I, \varepsilon)|$ iterations of the **for** loop, the running time of the algorithm is bounded by

$$|D| \cdot \tilde{O}(\ell^{(k-1)\cdot\ell} \cdot k) \cdot |I|^{O(1)} \leq (2\ell \cdot k \cdot \varepsilon^{-2} + 1) \cdot \tilde{O}(\ell^{(k-1)\cdot\ell} \cdot k) \cdot |I|^{O(1)} = \tilde{O}\left(\ell^{(k-1)\cdot\ell+1} \cdot k^2 \cdot \varepsilon^{-2}\right) \cdot |I|^{O(1)}.$$

The first inequality follows from (4) and (5). As in this case $\ell^{(k-1)\cdot\ell} \cdot k^2 \cdot \varepsilon^{-2} \leq |I|$, we have the desired running time.

For the cardinality of R, note that by Lemma 6 $\mathrm{OPT}(I) \geq \alpha \geq \frac{\mathrm{OPT}(I)}{2\ell}$. Thus, by Lemma 14, for all $r \in D$, $\mathsf{ExSet}(I, \varepsilon, \alpha, r)$ is an exchange set satisfying $|\mathsf{ExSet}(I, \varepsilon, \alpha, r)| = \tilde{O}(\ell^{(k-1)\cdot\ell} \cdot k)$. Then,

$$|R| \leq |D| \cdot \tilde{O}(\ell^{(k-1)\cdot\ell} \cdot k) \leq (2\ell \cdot k \cdot \varepsilon^{-2} + 1) \cdot \tilde{O}(\ell^{(k-1)\cdot\ell} \cdot k) = \tilde{O}\left(\ell^{(k-1)\cdot\ell+1} \cdot k^2 \cdot \varepsilon^{-2}\right).$$

The second inequality follows from (4) and (5).

To conclude, we show that R is a representative set. By Lemma 14, for all $r \in D$, it holds that $\mathsf{ExSet}(I, \varepsilon, \alpha, r)$ is an exchange set for I, ε, α, and r. Therefore, $R \cap \mathcal{K}_r(\alpha)$ is an exchange set for I, ε, α, for all $r \in D$. Hence, by Lemma 10, R is a representative set of I and ε. ◀

▶ **Theorem 3.** *There is an algorithm that given a* BM *instance* $I = (E, \mathcal{M}, c, p, B, k, \ell)$ *and* $0 < \varepsilon < \frac{1}{2}$*, returns in time* $|I|^{O(1)}$ *a representative set* $R \subseteq E$ *of* I *and* ε *such that* $|R| = \tilde{O}\left(\ell^{(k-1)\cdot\ell} \cdot k^2 \cdot \varepsilon^{-2}\right)$.

Proof of Theorem 3. The statement of the lemma follows from Lemma 15. ◀

3 An FPT Approximation Scheme

In this section we use the representative set constructed by Algorithm 1 to obtain an FPAS for BM. For the discussion below, fix a BM instance $I = (E, \mathcal{M}, c, p, B, k, \ell)$ and an error parameter $0 < \varepsilon < \frac{1}{2}$. Given the representative set R for I and ε output by algorithm RepSet, we derive an FPAS by exhaustive enumeration over all solutions of I within R. The pseudocode of our FPAS is given in Algorithm 2.

▶ **Lemma 16.** *Given a* BM *instance* $I = (E, \mathcal{M}, c, p, B, k, \ell)$ *and* $0 < \varepsilon < \frac{1}{2}$*, Algorithm 2 returns in time* $|I|^{O(1)} \cdot \tilde{O}\left(\ell^{k^2\cdot\ell} \cdot k^{2k} \cdot \varepsilon^{-2k}\right)$ *a solution for* I *of profit at least* $(1 - 2\varepsilon) \cdot \mathrm{OPT}(I)$.

We give the proof at the end of this section. We can now prove our main result.

▶ **Theorem 4.** *For any* BM *instance* $I = (E, \mathcal{M}, c, p, B, k, \ell)$ *and* $0 < \varepsilon < \frac{1}{2}$*, there is an* FPAS *whose running time is* $|I|^{O(1)} \cdot \tilde{O}\left(\ell^{k^2\cdot\ell} \cdot k^{O(k)} \cdot \varepsilon^{-2k}\right)$.

Algorithm 2 FPAS$(I = (E, \mathcal{M}, c, p, B, k, \ell), \varepsilon)$.

```
input  : A BM instance I and an error parameter 0 < ε < 1/2.
output : A solution for I.
1 Initialize an empty solution A ← ∅.
2 Construct R ← RepSet(I, ε).
3 for F ⊆ R s.t. |F| ≤ k and F is a solution of I do
4     if p(F) > p(A) then
5         Update A ← F
6 Return A.
```

Proof of Theorem 4. The statement of the lemma follows from Lemma 16 by using in Algorithm 2 an error parameter $\varepsilon' = \frac{\varepsilon}{2}$. ◀

For the proof of Lemma 16, we use the next auxiliary lemmas.

▶ **Lemma 17.** *Given a* BM *instance* $I = (E, \mathcal{M}, c, p, B, k, \ell)$ *and* $0 < \varepsilon < \frac{1}{2}$, *Algorithm 2 returns a solution for* I *of profit at least* $(1 - 2\varepsilon) \cdot \mathrm{OPT}(I)$.

Proof. By Lemma 15, it holds that $R = \mathsf{RepSet}(I, \varepsilon)$ is a representative set of I and ε. Therefore, by Definition 2, there is a solution S for I such that $S \subseteq R$, and

$$p(S) \geq (1 - 2\varepsilon) \cdot \mathrm{OPT}(I). \tag{9}$$

Since S is a solution for I, it follows that $S \in \mathcal{M}_k$ and therefore $|S| \leq k$. Thus, there is an iteration of Step 3 in which $F = S$, and therefore the set A returned by the algorithm satisfies $p(A) \geq p(S) \geq (1 - 2\varepsilon) \cdot \mathrm{OPT}(I)$. Also, the set A returned by the algorithm must be a solution for I: If $A = \emptyset$ the claim trivially follows since $\emptyset$ is a solution for I. Otherwise, the value of A has been updated in Step 5 of Algorithm 2 to be some set $F \subseteq R$, but this step is reached only if F is a solution for I. ◀

▶ **Lemma 18.** *Given a* BM *instance* $I = (E, \mathcal{M}, c, p, B, k, \ell)$ *and* $0 < \varepsilon < \frac{1}{2}$, *the running time of Algorithm 2 is* $|I|^{O(1)} \cdot \tilde{O}\left(\ell^{k^2 \cdot \ell} \cdot k^{2k} \cdot \varepsilon^{-2k}\right)$.

Proof. Let $W' = \left\{F \subseteq R \;\middle|\; F \in \mathcal{M}_k, c(F) \leq B\right\}$ be the solutions considered in Step 3 of Algorithm 2, and let $W = \left\{F \subseteq R \;\middle|\; |F| \leq k\right\}$. Observe that the number of iterations of Step 3 of Algorithm 2 is bounded by $|W|$, since $W' \subseteq W$ and for each $F \in W$ we can verify in polynomial time if $F \in W'$. Thus, it suffices to upper bound W.

By a simple counting argument, we have that

$$\begin{aligned} |W| \leq\ & (|R| + 1)^k \\ \leq\ & \tilde{O}\left(\left(\ell^{(k-1)\cdot\ell+1} \cdot k^2 \cdot \varepsilon^{-2}\right)^k\right) \\ =\ & \tilde{O}\left(\ell^{k^2\cdot\ell} \cdot k^{2k} \cdot \varepsilon^{-2k}\right) \end{aligned} \tag{10}$$

The first equality follows from Lemma 15. Hence, by (10), the number of iterations of the **for** loop in Step 3 is bounded by $\tilde{O}\left(\ell^{k^2\cdot\ell} \cdot k^{2k} \cdot \varepsilon^{-2k}\right)$. In addition, the running time of each iteration is at most $|I|^{O(1)}$. Finally, the running time of the steps outside the **for** loop is $|I|^{O(1)}$, by Lemma 15. Hence, the running time of Algorithm 2 can be bounded by $|I|^{O(1)} \cdot \tilde{O}\left(\ell^{k^2\cdot\ell} \cdot k^{2k} \cdot \varepsilon^{-2k}\right)$. ◀

Proof of Lemma 16. The proof follows from Lemmas 17 and 18. ◀

4 Hardness Results

In this section we prove Theorem 1 and Theorem 1. In the proof of Theorem 1, we use a reduction from the k-*subset sum (KSS)* problem. The input for KSS is a set $X = \{x_1, \ldots, x_n\}$ of strictly positive integers and two positive integers $T, k > 0$. We need to decide if there is a subset $S \subseteq [n], |S| = k$ such that $\sum_{i \in S} x_i = T$, where the problem is parameterized by k. KSS is known to be W[1]-hard [12].

▶ **Theorem 1.** BM *is* $W[1]$-hard*.*

Proof of Theorem 1. Let U be a KSS instance with the set of numbers $E = [n]$, target value T, and k. We define the following BM instance $I = (E, \mathcal{M}, c, p, B, k, \ell)$.

1. $\mathcal{M}$ is a 1-matchoid $\mathcal{M} = \{(E, \mathcal{I})\}$ such that $\mathcal{I} = 2^E$. That is, $\mathcal{M}$ is a single uniform matroid whose independent sets are all possible subsets of E.
2. For any $i \in E = [n]$ define $c(i) = p(i) = x_i + 2 \cdot \sum_{j \in [n]} x_j$.
3. Define the budget as $B = T + 2k \cdot \sum_{j \in [n]} x_j$.

▷ Claim 19. If there is a solution for U then there is a solution for I of profit B.

Proof. Let $S \subseteq [n], |S| = k$ such that $\sum_{i \in S} x_i = T$. Then,

$$c(S) = p(S) = \sum_{i \in S} \left(x_i + 2 \cdot \sum_{j \in [n]} x_j \right) = T + |S| \cdot 2 \cdot \sum_{j \in [n]} x_j = T + 2k \cdot \sum_{j \in [n]} x_j = B.$$

By the above, and as $S \in \mathcal{M}_k$, S is also a solution for I of profit exactly B. ◁

▷ Claim 20. If there is a solution for I of profit at least B then there is a solution for U.

Proof. Let F be a solution for I of profit at least B. Then, $p(F) = c(F) \leq B$, since F satisfies the budget constraint. As $p(F) \geq B$, we conclude that

$$p(F) = c(F) = B. \tag{11}$$

We now show that F is also a solution for U. First, assume towards contradiction that $|F| \neq k$. If $|F| < k$ then

$$p(F) = \sum_{i \in F} x_i + |F| \cdot 2 \cdot \sum_{j \in [n]} x_j \leq \sum_{i \in F} x_i + (k-1) \cdot 2 \cdot \sum_{j \in [n]} x_j \leq 2k \cdot \sum_{j \in [n]} x_j < B.$$

We reach a contradiction to (11). Since F is a solution for I it holds that $F \in \mathcal{M}_k$; thus, $|F| \leq k$. By the above, $|F| = k$. Therefore,

$$\sum_{i \in F} x_i = c(F) - |F| \cdot 2 \cdot \sum_{j \in [n]} x_j = c(F) - 2k \cdot \sum_{j \in [n]} x_j = B - 2k \cdot \sum_{j \in [n]} x_j = T. \quad ◁$$

By Claims 19 and 20, there is a solution for U if and only if there is a solution for I of profit at least B. Furthermore, the construction of I can be done in polynomial time in the encoding size of U. Hence, an FPT algorithm which finds an optimal solution for I can decide the instance U in FPT time. As KSS is known to be W[1]-hard [12], we conclude that BM is also W[1]-hard. ◀

In the proof of Theorem 5 we use a lower bound on the kernel size of *Perfect ℓ-Dimensional Matching* (ℓ-PDM), due to Dell and Marx [5, 6]. The input for the problem consists of the finite sets $U_1, \ldots U_\ell$ and $E \subseteq U_1 \times \ldots \times U_\ell$. Also, we have an ℓ-dimensional matching constraint $(E, \mathcal{I})$ to which we refer as the associated set system of the instance (i.e., $\mathcal{I}$ contains all subsets $S \subseteq E$ such that for any two distinct tuples $(e_1, \ldots, e_\ell), (f_1, \ldots, f_\ell) \in S$ and every $i \in [\ell]$ it holds that $e_i \neq f_i$). The instance is associated also with the parameter $k = \frac{n}{\ell}$, where $n = \sum_{j=1}^{\ell} |U_\ell|$. We refer to $|E|$ as the *number of tuples* in the instance. The objective is to find $S \in \mathcal{I}$ such that $|S| = k$. Let $J = (U_1, \ldots, U_\ell, E)$ denote an instance of ℓ-PDM We say J is a "yes" instance if such a set S exists; otherwise, J is a "no" instance. Observe that the parameter k is set such that if $S \in \mathcal{I}$ and $|S| = k$ then every element in $U_1 \cup \ldots \cup U_\ell$ appears in exactly one of the tuples in S.

▶ **Lemma 21** (Theorem 1.2 cf. [6]). *Let $\ell \geq 3$ and $\varepsilon > 0$. If* coNP $\not\subseteq$ NP/poly *then* ℓ-PDM *does not have a kernel in which the number of tuples is $O(k^{\ell-\varepsilon})$.*

▶ **Theorem 5.** *For any function $f : \mathbb{N} \to \mathbb{N}$, and $c_1, c_2 \in \mathbb{R}$ such that $c_2 - c_1 < 0$, there is no algorithm which finds for a given* BM *instance $I = (E, \mathcal{M}, c, p, B, k, \ell)$ and $0 < \varepsilon < \frac{1}{2}$ a representative set of size $O\left(f(\ell) \cdot k^{\ell - c_1} \cdot \frac{1}{\varepsilon^{c_2}}\right)$ of I and ε in time $|I|^{O(1)}$, unless* coNP $\subseteq$ NP/poly.

Proof of Theorem 5. Assume coNP $\not\subseteq$ NP/poly. Furthermore, assume towards a contradiction that there is a function $f : \mathbb{N} \to \mathbb{N}$, constants c_1, c_2, where $c_2 - c_1 < 0$, and an algorithm $\mathcal{A}$ that, given a BM instance $I = (E, \mathcal{M}, c, p, B, k, \ell)$ and $0 < \varepsilon < \frac{1}{2}$, finds in time $|I|^{O(1)}$ a representative set of I and ε of size $O\left(f(\ell) \cdot k^{\ell - c_1} \cdot \frac{1}{\varepsilon^{c_2}}\right)$. We use $\mathcal{A}$ to construct a kernel for 3-PDM.

Consider the following kernelization algorithm for 3-PDM. Let $J = (U_1, U_2, U_3, E)$ be the 3-PDM input instance. Define $n = |U_1| + |U_2| + |U_3|$, $\ell = 3$, and $k = \frac{n}{\ell}$. Furthermore, let $(E, \mathcal{I})$ be the set system associated with the instance, and let $\mathcal{M}$ be an ℓ-matchoid representing the set system $(E, \mathcal{I})$. Run $\mathcal{A}$ on the BM instance $I = (E, \mathcal{M}, , p, B, k, \ell)$ with $\varepsilon = \frac{1}{3k}$, where $c(e) = p(e) = 1$ for all $e \in E$ and $B = k$. Let $R \subseteq E$ be the output of $\mathcal{A}$. Return the 3-PDM instance $J' = (U_1, U_2, U_3, R)$.

Since $\mathcal{A}$ runs in polynomial time, the above algorithm runs in polynomial time as well. Moreover, as $k = \frac{n}{3}$ and $R \subseteq E$, it follows that the returned instance can be encoded using $O(k^4)$ bits. Let $(R, \mathcal{I}')$ be the set system associated with J'. Since $R \subseteq E$, it follows that $\mathcal{I}' \subseteq \mathcal{I}$. Hence, if there is $S \in \mathcal{I}'$ such that $|S| = k$, then $S \in \mathcal{I}$ as well. That is, if J' is a "yes" instance, so is J.

For the other direction, assume that J is a "yes" instance. That is, there is $S \in \mathcal{I}$ such that $|S| = k$. Then S is a solution for the BM instance I (observe that $c(S) = |S| = k = B$). Therefore, as R is a representative set of I and $\varepsilon = \frac{1}{3k}$, there is a solution T for I such that $T \subseteq R$, and

$$p(T) \geq (1 - 2\varepsilon) \cdot \mathrm{OPT}(I) \geq (1 - 2\varepsilon) \cdot p(S) = \left(1 - \frac{2}{3k}\right) \cdot p(S) = \left(1 - \frac{2}{3k}\right) \cdot k = k - \frac{2}{3}.$$

Since the profits are integral we have that $|T| = p(T) \geq k$. Furthermore $|T| \leq k$ (since T is a solution for I), and thus $|T| = k$. Since $T \in \mathcal{I}$ (as T is a solution for I) and $T \subseteq R$, it trivially holds that $T \in \mathcal{I}'$. That is, $T \in \mathcal{I}'$ and $|T| = k$. Hence, J' is a "yes" instance. We have showed that the above procedure is indeed a kernelization for 3-PDM.

Now, consider the size of R. Since $\mathcal{A}$ returns a representative set of size $O\left(f(\ell) \cdot k^{\ell - c_1} \cdot \frac{1}{\varepsilon^{c_2}}\right)$ it follows that

$$|R| = O\left(f(3) \cdot k^{3 - c_1} \cdot (3k)^{c_2}\right) = O\left(k^{3 - c_1 + c_2}\right).$$

As $c_2 - c_1 < 0$, we have a contradiction to Lemma 21. Thus, for any function $f : \mathbb{N} \to \mathbb{N}$ and constants c_1, c_2 satisfying $c_2 - c_1 < 0$, there is no algorithm which finds for a given BM instance $I = (E, \mathcal{M}, c, p, B, k, \ell)$ and $0 < \varepsilon < \frac{1}{2}$ a representative set of I and ε of size $O\left(f(\ell) \cdot k^{\ell - c_1} \cdot \frac{1}{\varepsilon^{c_2}}\right)$ in time $|I|^{O(1)}$. ◀

5 A Polynomial-time $\frac{1}{2 \cdot \ell}$-Approximation for BM

In this section we prove Lemma 6. The proof combines an existing approximation algorithm for the unbudgeted version of BM [23, 25] with the Lagrangian relaxation technique of [29]. As the results in [23, 25] are presented in the context of *ℓ-extendible* set systems, we first define these systems and use a simple argument to show that such systems are generalizations of matchoids. We refer the reader to [13] for further details about ℓ-extendible systems.

▶ **Definition 22.** *Given a finite set E, $\mathcal{I} \subseteq 2^E$, and $\ell \in \mathbb{N}$, we say that $(E, \mathcal{I})$ is an ℓ-*extendible system *if for every $S \in \mathcal{I}$ and $e \in E \setminus S$ there is $T \subseteq S$, where $|T| \leq \ell$, such that $(S \setminus T) \cup \{\ell\} \in \mathcal{I}$.*

The next lemma shows that an ℓ-matchoid is in fact an ℓ-extendible set system.

▶ **Lemma 23.** *For any $\ell \in \mathbb{N}_{>0}$ and an ℓ-Matchoid $\mathcal{M} = \{M_i = (E_i, \mathcal{I}_i)\}_{i \in [s]}$ on a set E, it holds that $(E, \mathcal{I}(\mathcal{M}))$ is an ℓ-extendible set system.*

Proof. Let $S \in \mathcal{I}(\mathcal{M})$ and $e \in E \setminus S$. As $\mathcal{M}$ is an ℓ-matchoid, there is $H \subseteq [s]$ of cardinality $|H| \leq \ell$ such that for all $i \in [s] \setminus H$ it holds that $e \notin E_i$ and for all $i \in H$ it holds that $e \in E_i$. Since for all $i \in H$ it holds that $(E_i, \mathcal{I}_i)$ is a matroid, either $(S \cap E_i) \cup \{e\} \in \mathcal{I}_i$, or there is $a_i \in S \cap E_i$ such that $((S \cap E_i) \setminus \{a_i\}) \cup \{e\} \in \mathcal{I}_i$ (this follows by repeatedly adding elements from $S \cap E_i$ to $\{e\}$ using the exchange property of the matroid $(E_i, \mathcal{I}_i)$). Let $L = \{i \in H \mid (S \cap E_i) \cup \{e\} \notin \mathcal{I}_i\}$. Then, there are $|L|$ elements $T = \{a_i\}_{i \in L}$ such that for all $i \in L$ it holds that $((S \cap E_i) \setminus \{a_i\}) \cup \{e\} \in \mathcal{I}_i$ and for all $i \in H \setminus L$ it holds that $(S \cap E_i) \cup \{e\} \in \mathcal{I}_i$. Thus, it follows that $(S \setminus T) \cup \{e\} \in \mathcal{I}(\mathcal{M})$ by the definition of a matchoid. Since $|T| = |L| \leq |H| \leq \ell$, we have the statement of the lemma. ◀

Proof of Lemma 6. Consider the BM problem with no budget constraint (equivalently, $c(E) \leq B$) that we call the *maximum weight matchoid maximization (MWM)* problem. By Lemma 23, MWM is a special case of the *maximum weight ℓ-extendible system maximization* problem, which admits $\frac{1}{\ell}$-approximation [23, 25].[5] Therefore, using Theorem 3.1 in [29], we have the following. There is an algorithm that, given some $\varepsilon > 0$, returns a solution for the BM instance I of profit at least $\left(\frac{\frac{1}{\ell}}{\frac{1}{\ell}+1} - \varepsilon\right) \cdot \text{OPT}(I)$, and whose running time is $|I|^{O(1)} \cdot O(\log(\varepsilon^{-1}))$. Now, we can set $\varepsilon = \frac{\frac{1}{\ell}}{\frac{1}{\ell}+1} - \frac{1}{2\ell}$; then, the above algorithm has a running time $|I|^{O(1)}$, since ε^{-1} is polynomial in ℓ and $\ell \leq |I|$. Moreover, the algorithm returns a solution S for I, such that

$$\text{OPT}(I) \geq p(S) \geq \left(\frac{\frac{1}{\ell}}{\frac{1}{\ell}+1} - \varepsilon\right) \cdot \text{OPT}(I) = \frac{1}{2\ell} \cdot \text{OPT}(I).$$

To conclude, we define the algorithm ApproxBM which returns $\alpha = p(S)$. By the above discussion, $\text{OPT}(I) \geq \alpha \geq \frac{\text{OPT}(I)}{2\ell}$, and the running time of ApproxBM is $|I|^{O(1)}$. ◀

[5] The algorithm of [23] can be applied also in the more general setting of ℓ-systems. For more details on such set systems, see, e.g., [13].

6 Discussion

In this paper we present an FPT-approximation scheme (FPAS) for the budgeted ℓ-matchoid problem (BM). As special cases, this yields FPAS for the budgeted ℓ-dimensional matching problem (BDM) and the budgeted ℓ-matroid intersection problem (BMI). While the unbudgeted version of BM has been studied earlier from parameterized viewpoint, the budgeted version is studied here for the first time.

We show that BM parameterized by the solution size is $W[1]$-hard already with a degenerate matroid constraint (Theorem 1); thus, an exact FPT time algorithm is unlikely to exist. Furthermore, the special case of unbudgeted ℓ-dimensional matching problem is APX-hard, already for $\ell = 3$, implying that PTAS for this problem is also unlikely to exist. These hardness results motivated the development of an FPT-approximation scheme for BM.

Our FPAS relies on the notion of representative set – a small cardinality subset of the ground set of the original instance which preserves the optimum value up to a small factor. We note that representative sets are not *lossy kernels* [31] as BM is defined in an oracle model; thus, the definitions of kernels or lossy kernels do not apply to our problem. Nevertheless, for some variants of BM in which the input is given explicitly (for instance, this is possible for BDM) our construction of representative sets can be used to obtain an approximate kernelization scheme.

Our results also include a lower bound on the minimum possible size of a representative set for BM which can be computed in polynomial time (Theorem 5). The lower bound is based on the special case of the budgeted ℓ-dimensional matching problem (BDM). We note that there is a significant gap between the size of the representative sets found in this paper and the lower bound. This suggests the following questions for future work.

- Is there a representative set for the special case of BDM whose size matches the lower bound given in Theorem 5?
- Can the generic structure of ℓ-matchoids be used to derive an improved lower bound on the size of a representative set for general BM instances?

The budgeted ℓ-matchoid problem can be naturally generalized to the d-budgeted ℓ-matchoid problem (d-BM). In the d-budgeted version, both the costs and the budget are replaced by d-dimensional vectors, for some constant $d \geq 2$. A subset of elements is feasible if it is an independent set of the ℓ-matchoid, and the total cost of the elements in each dimension is bounded by the budget in this dimension. The problem is a generalization of the d-dimensional knapsack problem (d-KP), the special case of d-BM in which the feasible sets of the matchoid are all subsets of E. A PTAS for d-KP was first given in [17], and the existence of an *efficient* polynomial time approximation scheme was ruled out in [28]. PTASs for the special cases of d-BM in which the matchoid is a single matroid, matroid intersection or a matching constraint were given in [3, 20]. It is likely that the lower bound in [28] can be used also to rule out the existence of an FPAS for d-BM. However, the question whether d-BM admits a $(1-\varepsilon)$-approximation in time $O\left(f(k+\ell) \cdot n^{g(\varepsilon)}\right)$, for some functions f and g, remains open.

References

1 André Berger, Vincenzo Bonifaci, Fabrizio Grandoni, and Guido Schäfer. Budgeted matching and budgeted matroid intersection via the gasoline puzzle. *Mathematical Programming*, 128(1):355–372, 2011.

2 Andreas Björklund, Thore Husfeldt, Petteri Kaski, and Mikko Koivisto. Narrow sieves for parameterized paths and packings. *Journal of Computer and System Sciences*, 87:119–139, 2017.

3 Chandra Chekuri, Jan Vondrák, and Rico Zenklusen. Multi-budgeted matchings and matroid intersection via dependent rounding. In *Proceedings of the twenty-second annual ACM-SIAM symposium on Discrete Algorithms*, pages 1080–1097. SIAM, 2011.

4 Jianer Chen, Qilong Feng, Yang Liu, Songjian Lu, and Jianxin Wang. Improved deterministic algorithms for weighted matching and packing problems. *Theoretical computer science*, 412(23):2503–2512, 2011.

5 Holger Dell and Dániel Marx. Kernelization of packing problems. In *Proceedings of the twenty-third annual ACM-SIAM symposium on Discrete Algorithms*, pages 68–81. SIAM, 2012.

6 Holger Dell and Dániel Marx. Kernelization of packing problems. *arXiv preprint arXiv:1812.03155*, 2018.

7 Ilan Doron-Arad, Ariel Kulik, and Hadas Shachnai. Budgeted matroid maximization: a parameterized viewpoint. *arXiv preprint arXiv:2307.04173*, 2023.

8 Ilan Doron-Arad, Ariel Kulik, and Hadas Shachnai. An EPTAS for budgeted matching and budgeted matroid intersection via representative sets. In *50th International Colloquium on Automata, Languages, and Programming (ICALP)*, 2023.

9 Ilan Doron-Arad, Ariel Kulik, and Hadas Shachnai. An EPTAS for budgeted matroid independent set. In *Symposium on Simplicity in Algorithms (SOSA)*, pages 69–83, 2023.

10 Ilan Doron-Arad, Ariel Kulik, and Hadas Shachnai. An FPTAS for budgeted laminar matroid independent set. *Operations Research Letters*, 51(6):632–637, 2023. `doi:10.1016/j.orl.2023.10.005`.

11 Ilan Doron-Arad, Ariel Kulik, and Hadas Shachnai. Tight lower bounds for weighted matroid problems. *arXiv preprint arXiv:2307.07773*, 2023.

12 Rod G Downey and Michael R Fellows. Fixed-parameter tractability and completeness ii: On completeness for W[1]. *Theoretical Computer Science*, 141(1-2):109–131, 1995.

13 Moran Feldman, Joseph Naor, Roy Schwartz, and Justin Ward. Improved approximations for k-exchange systems. In *Algorithms–ESA 2011: 19th Annual European Symposium, Saarbrücken, Germany, September 5-9, 2011. Proceedings 19*, pages 784–798. Springer, 2011.

14 Andreas Emil Feldmann, Euiwoong Lee, and Pasin Manurangsi. A survey on approximation in parameterized complexity: Hardness and algorithms. *Algorithms*, 13(6):146, 2020.

15 Fedor V Fomin, Daniel Lokshtanov, Fahad Panolan, and Saket Saurabh. Efficient computation of representative families with applications in parameterized and exact algorithms. *Journal of the ACM (JACM)*, 63(4):1–60, 2016.

16 Fedor V Fomin, Daniel Lokshtanov, Saket Saurabh, and Meirav Zehavi. *Kernelization: theory of parameterized preprocessing*. Cambridge University Press, 2019.

17 Alan M Frieze, Michael RB Clarke, et al. Approximation algorithms for the m-dimensional 0-1 knapsack problem: worst-case and probabilistic analyses. *European Journal of Operational Research*, 15(1):100–109, 1984.

18 Michael R Garey and David S Johnson. *Computers and intractability*, volume 174. freeman San Francisco, 1979.

19 Prachi Goyal, Neeldhara Misra, Fahad Panolan, and Meirav Zehavi. Deterministic algorithms for matching and packing problems based on representative sets. *SIAM Journal on Discrete Mathematics*, 29(4):1815–1836, 2015.

20 Fabrizio Grandoni and Rico Zenklusen. Approximation schemes for multi-budgeted independence systems. In *European Symposium on Algorithms*, pages 536–548. Springer, 2010.

21 Chien-Chung Huang and Justin Ward. FPT-algorithms for the ℓ-matchoid problem with a coverage objective. *arXiv preprint arXiv:2011.06268*, 2020.

22 Chien-Chung Huang and Justin Ward. FPT-algorithms for the ℓ-matchoid problem with a coverage objective. *SIAM Journal on Discrete Mathematics*, 2023.

23 Th Jenkyns. The efficacy of the "greedy" algorithm. In *Proc. 7th Southeastern Conf. on Combinatorics, Graph Theory and Computing*, pages 341–350, 1976.

24 Per M Jensen and Bernhard Korte. Complexity of matroid property algorithms. *SIAM Journal on Computing*, 11(1):184–190, 1982.

25 Stasys Jukna. *Extremal combinatorics: with applications in computer science*, volume 571. Springer, 2011.

26 Viggo Kann. Maximum bounded 3-dimensional matching is max snp-complete. *Information Processing Letters*, 37(1):27–35, 1991.

27 Ioannis Koutis and Ryan Williams. Limits and applications of group algebras for parameterized problems. In *Automata, Languages and Programming: 36th International Colloquium, ICALP 2009, Rhodes, Greece, July 5-12, 2009, Proceedings, Part I 36*, pages 653–664. Springer, 2009.

28 Ariel Kulik and Hadas Shachnai. There is no EPTAS for two-dimensional knapsack. *Information Processing Letters*, 110(16):707–710, 2010.

29 Ariel Kulik, Hadas Shachnai, and Gal Tamir. On lagrangian relaxation for constrained maximization and reoptimization problems. *Discrete Applied Mathematics*, 296:164–178, 2021.

30 Daniel Lokshtanov, Pranabendu Misra, Fahad Panolan, and Saket Saurabh. Deterministic truncation of linear matroids. *ACM Transactions on Algorithms (TALG)*, 14(2):1–20, 2018.

31 Daniel Lokshtanov, Fahad Panolan, MS Ramanujan, and Saket Saurabh. Lossy kernelization. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 224–237, 2017.

32 László Lovász. Matroid matching and some applications. *Journal of Combinatorial Theory, Series B*, 28(2):208–236, 1980.

33 Silvano Martello and Paolo Toth. *Knapsack problems: algorithms and computer implementations.* John Wiley & Sons, Inc., 1990.

34 Dániel Marx. Parameterized complexity and approximation algorithms. *The Computer Journal*, 51(1):60–78, 2008.

35 Dániel Marx. A parameterized view on matroid optimization problems. *Theoretical Computer Science*, 410(44):4471–4479, 2009.

36 Ram Ravi and Michel X Goemans. The constrained minimum spanning tree problem. In *Scandinavian Workshop on Algorithm Theory*, pages 66–75. Springer, 1996.

37 Alexander Schrijver et al. *Combinatorial optimization: polyhedra and efficiency*, volume 24. Springer, 2003.