Randomized Strategies for Cardinality Robustness in the Knapsack Problem☆

Yusuke Kobayashi[1]

University of Tsukuba, Tsukuba, Ibaraki 305-8573, Japan

Kenjiro Takazawa[2]

Hosei University, Tokyo 184-8584, Japan

Abstract

We consider the following zero-sum game related to the knapsack problem. Given an instance of the knapsack problem, Alice chooses a knapsack solution and Bob, knowing Alice's solution, chooses a cardinality k. Then, Alice obtains a payoff equal to the ratio of the profit of the best k items in her solution to that of the best solution of size at most k. For $\alpha > 0$, a knapsack solution is called *α-robust* if it guarantees payoff α. If Alice adopts a deterministic strategy, the objective of Alice is to find a max-robust knapsack solution. By applying the argument in Kakimura and Makino (2013) for robustness in general independence systems, a $(1/\sqrt{\mu})$-robust solution exists and is found in polynomial time, where μ is the exchangeability of the independence system.

In the present paper, we address randomized strategies for this zero-sum game. Randomized strategies in robust independence systems are introduced by Matuschke, Skutella, and Soto (2015) and they presented a randomized strategy with $(1/\ln 4)$-robustness for a certain class of independence systems. The knapsack problem, however, does not belong to this class. We first establish the intractability of the knapsack problem by showing an instance such that the robustness of an arbitrary randomized strategy is both $\mathrm{O}((\log\log\mu)/\log\mu)$ and $\mathrm{O}((\log\log\rho)/\log\rho)$, where $\rho := \frac{\text{(the size of a maximum feasible set)}}{\text{(the size of a minimum infeasible set)}-1}$. We then exhibit the power of randomness by designing two randomized strategies with robustness $\Omega(1/\log\mu)$ and $\Omega(1/\log\rho)$, respectively, which substantially improve upon that of known deterministic strategies and almost attain the above upper bounds. It is also noteworthy that our strategy applies to not only the knapsack problem but also independence systems for which an (approximately) optimal

☆A preliminary version of this paper appears in ANALCO 2016 [8].

Email addresses: kobayashi@sk.tsukuba.ac.jp (Yusuke Kobayashi), takazawa@hosei.ac.jp (Kenjiro Takazawa)

[1]Supported by JST, ERATO, Kawarabayashi Project, by JSPS KAKENHI Grant Number 24700004, and by MEXT KAKENHI Grant Number 24106002.

[2]Partially supported by JSPS/MEXT KAKENHI Grant Numbers 26280001 and 16K16012.

solution under a cardinality constraint is computable.

Keywords: Robust independence system, Randomized strategy, Knapsack problem, Exchangeability
2010 MSC: 90C27, 68W20, 68Q25

1. Introduction

1.1. Cardinality robustness in independence systems

Cardinality robustness in independence systems is introduced by Hassin and Rubinstein [4], defined as follows. Let $(E, \mathcal{F})$ be an independence system. That is, E is a finite set of items and $\mathcal{F} \subseteq 2^E$ is the *feasible set family* satisfying that $\emptyset \in \mathcal{F}$ and $X \subseteq Y \in \mathcal{F}$ implies $X \in \mathcal{F}$. A feasible set is often referred to as a *solution*. Let $p_e \in \mathbf{R}_+$ represent the profit of item $e \in E$, and let $\mathrm{OPT}_k \subseteq E$ be a feasible set maximizing its profit among those of size at most k. That is, OPT_k satisfies that $\mathrm{OPT}_k \in \mathcal{F}$, $|\mathrm{OPT}_k| \leq k$, and $p(\mathrm{OPT}_k) = \max\{p(X) \mid X \in \mathcal{F}, |X| \leq k\}$, where the profit $p(X)$ of a feasible set X is defined by $p(X) := \sum_{e \in X} p_e$. For $X \in \mathcal{F}$, let $X(k)$ denote a subset of X satisfying that $|X(k)| \leq k$ and $p(X(k)) = \max\{p(X') \mid X' \subseteq X, |X'| \leq k\}$. Intuitively, $X(k)$ consists of k p-highest items in X. For $\alpha > 0$, a feasible set $X \in \mathcal{F}$ is called *α-robust* if $p(X(k)) \geq \alpha \cdot p(\mathrm{OPT}_k)$ for every positive integer k.

Our problem is to find a feasible set with large robustness. This is described as the following zero-sum game.

> Alice chooses a feasible set $X \in \mathcal{F}$, and Bob, knowing Alice's set, chooses a cardinality k. Then, Alice obtains a payoff $p(X(k))/p(\mathrm{OPT}_k)$.

In this zero-sum game, the objective of Alice is to find a feasible set with maximum robustness.

It is not difficult to see that, if $\mathcal{F}$ is the independent set family of a matroid on E, then a greedy solution[3] is 1-robust. More generally, Hassin and Rubinstein [4] proved that a greedy solution is $r(\mathcal{F})$-robust, where $r(\mathcal{F})$ is the *rank quotient* of $(E, \mathcal{F})$ [5, 9].

A p^2-optimal solution, i.e., a feasible set $X \in \mathcal{F}$ maximizing $\sum_{e \in X} p_e^2$, often has larger robustness than a greedy solution. Hassin and Rubinstein [4] showed that a p^2-optimal matching is $(1/\sqrt{2})$-robust, and there exist graphs not containing an α-robust matching for an arbitrary $\alpha > 1/\sqrt{2}$. Fujita, Kobayashi, and Makino [3] discussed the case where $\mathcal{F}$ is defined by matroid intersection, i.e., common independent sets of two matroids on E, and proved that a p^2-optimal common independent set is $(1/\sqrt{2})$-robust. It is also shown in [3] that determining whether a graph has an α-robust matching is NP-hard for an arbitrary

[3]The *greedy algorithm* for an independence system is defined as follows. Sort the elements $e \in E$ by profit p_e, i.e., $E = \{e_1, \ldots, e_n\}$ and $p_{e_1} \geq \cdots \geq p_{e_n}$. Let $X_0 = \emptyset$, and for $i = 1, \ldots, n$, let $X_i = X_{i-1} \cup \{e_i\}$ if $X_{i-1} \cup \{e_i\} \in \mathcal{F}$, and $X_i = X_{i-1}$ otherwise. The algorithm returns X_n, called a *greedy solution*.

$\alpha > 1/\sqrt{2}$. Analysis for general independence systems is due to Kakimura and Makino [6], who proved that a p^2-optimal feasible set is a $(1/\sqrt{\mu(\mathcal{F})})$-robust solution, where $\mu(\mathcal{F})$, the *exchangeability* of $(E, \mathcal{F})$, is defined as the minimum integer μ satisfying that

$$\forall X, Y \in \mathcal{F},\ \forall e \in Y - X,\ \exists Z \subseteq X - Y \text{ s.t. } |Z| \leq \mu,\ (X - Z) \cup \{e\} \in \mathcal{F}. \quad (1)$$

In [6], it is also shown that the above robustness is tight in the sense that for an arbitrary positive integer μ, there exists an independence system $(E, \mathcal{F})$ such that $\mu(\mathcal{F}) = \mu$ and no α-robust solution exists for arbitrary $\alpha > 1/\sqrt{\mu}$.

Kakimura, Makino, and Seimi [7] focused on the case where $(E, \mathcal{F})$ is defined by an instance of the knapsack problem. An instance (E, p, w, C) of the knapsack problem consists of the set E of items, the profit vector $p \in \mathbf{R}_+^E$, the weight vector $w \in \mathbf{R}_+^E$, and the capacity $C \in \mathbf{R}_+$. A subset $X \subseteq E$ is feasible if its weight $w(X) := \sum_{e \in X} w_e$ is at most the capacity, i.e., $\mathcal{F} = \{X \subseteq E \mid w(X) \leq C\}$. Kakimura, Makino, and Seimi [7] proved that the problem of computing a knapsack solution with the maximum robustness is weakly NP-hard, and also presented a fully polynomial-time approximation scheme (FPTAS) for this problem.

A similar problem is studied in [2], in which we choose a permutation of the items to be packed without knowing the capacity.

1.2. Randomized strategies

The above results correspond to deterministic strategies (or pure strategies) of the zero-sum game. Matuschke, Skutella, and Soto [11] introduced randomized strategies (or mixed strategies) for the robust independence systems. In this setting, Alice calls a probability distribution on the feasible sets, and Bob, knowing the distribution of Alice, chooses an integer k. The robustness of Alice's strategy is defined by the expected payoff. That is, if Alice chooses a distribution in which a solution X_i has probability λ_i, then the robustness of this strategy is

$$\min_k \frac{\mathbf{E}[p(X_i(k))]}{p(\mathrm{OPT}_k)} = \min_k \frac{\sum_i \lambda_i p(X_i(k))}{p(\mathrm{OPT}_k)}.$$

For the robust matching case, Matuschke, Skutella, and Soto [11] presented a randomized strategy with robustness $1/\ln 4$, which shows that randomized strategies break the bound on the robustness $1/\sqrt{2}$ of the deterministic strategies. They further showed that this strategy attains robustness $1/\ln 4$ for *bit-concave* independence systems. Examples of bit-concave independence systems include matroid intersection, b-matchings, strongly base orderable matchoids, strongly base orderable matroid parity systems.

A related problem is studied in [10], in which Alice chooses a randomized strategy and Bob chooses a cost function to maximize the regret of Alice.

1.3. Our results

We address randomized strategies for the robust independence systems defined by an instance of the knapsack problem. Note that those independence systems are not necessarily bit-concave, and hence the method in [11] cannot be applied. In what follows, we assume that $\mathcal{F} \neq 2^E$ and $\{e\} \in \mathcal{F}$ for every $e \in E$. That is, $w_e \leq C$ for every $e \in E$ and $w(E) > C$.

We provide upper and lower bounds for the robustness in terms of the exchangeability $\mu(\mathcal{F})$ and a new parameter $\rho(\mathcal{F})$, defined by

$$\rho(\mathcal{F}) := \frac{a_{\max}}{a_{\min}}, \quad a_{\max} := \max\{|X| \mid X \in \mathcal{F}\}, \quad a_{\min} := \min\{|X| \mid X \notin \mathcal{F}\} - 1. \tag{2}$$

We remark that the parameters $\mu(\mathcal{F})$ and $\rho(\mathcal{F})$ represent the intractability of the independence system $(E, \mathcal{F})$. Clearly $\mu(\mathcal{F}) \geq 1$ and $\rho(\mathcal{F}) \geq 1$ hold. Moreover, $\mu(\mathcal{F}) = 1$ holds if and only if $\mathcal{F}$ is the independent set family of a matroid, and $\rho(\mathcal{F}) = 1$ holds if and only if $\mathcal{F}$ is the independent set family of a uniform matroid. If $\mathcal{F}$ is defined by the matchings in a graph, then $\mu(\mathcal{F}) \leq 2$. For the problem of finding a feasible set X maximizing $p(X)$, the greedy algorithm attains a $(1/\mu(\mathcal{F}))$-approximation [12]. Further, we show that a greedy solution yields $(1/\rho(\mathcal{F}))$-approximation as well (see Proposition 3), and construct examples for which the greedy algorithm has approximation ratio no better than $1/\mu(\mathcal{F})$ and $1/\rho(\mathcal{F})$ (see § 2). We also note that $\rho(\mathcal{F})$ is a parameter whose definition is similar to $1/r(\mathcal{F})$. Thus, roughly speaking, the larger $\mu(\mathcal{F})$ or $\rho(\mathcal{F})$ becomes, the harder optimization over $(E, \mathcal{F})$ becomes.

We establish the intractability of the robust knapsack problem by showing a family of instances which do not admit a randomized strategy with constant robustness. Indeed, for those instances, we prove that the robustness of an arbitrary randomized strategy is both $\mathrm{O}((\log\log\mu(\mathcal{F}))/\log\mu(\mathcal{F}))$ and $\mathrm{O}((\log\log\rho(\mathcal{F}))/\log\rho(\mathcal{F}))$.

We then exhibit the power of randomness by designing two randomized strategies with robustness $\Omega(1/\log\mu(\mathcal{F}))$ and $\Omega(1/\log\rho(\mathcal{F}))$, respectively. These lower bounds substantially improve upon that of known deterministic strategies, and almost attain the above upper bounds. Roughly speaking, the $\Omega(1/\log\rho(\mathcal{F}))$-robust strategy is a uniform distribution of the optimal solutions under different cardinality constraints, which are efficiently computed by an FPTAS [1]. In the $\Omega(1/\log\mu(\mathcal{F}))$-robust strategy, we modify the $\Omega(1/\log\rho(\mathcal{F}))$-robust strategy so that some items in the optimal solution are always chosen, which helps attaining good robustness when $\mu(\mathcal{F})$ is small.

Furthermore, we extend the aforementioned results to general independence systems. The $\Omega(1/\log\rho(\mathcal{F}))$-robust strategy can be applied to general independence systems. We provide upper bounds $\mathrm{O}(1/\log\rho(\mathcal{F}))$ and $\mathrm{O}(1/\log\mu(\mathcal{F}))$ on robustness, which prove the tightness of our $\Omega(1/\log\rho(\mathcal{F}))$-robust strategy.

1.4. Organization of the paper

The rest of this paper is organized as follows. In § 2, we show an instance of the knapsack problem for which no randomized strategy attains con-

Table 1: An instance denying a constant robustness. The capacity is $C = M^{2T}$.

type	w	p	number of items	density p/w	total profit
0	M^{2T}	M^{2T}	1	1	M^{2T}
1	M^{2T-2}	M^{2T-1}	M^2	M	M^{2T+1}
2	M^{2T-4}	M^{2T-2}	M^4	M^2	M^{2T+2}
		$\vdots$			
i	M^{2T-2i}	M^{2T-i}	M^{2i}	M^i	M^{2T+i}
		$\vdots$			
$T-1$	M^2	M^{T+1}	M^{2T-2}	M^{T-1}	M^{3T-1}
T	1	M^T	M^{2T}	M^T	M^{3T}

stant robustness, and provide upper bounds $\mathrm{O}((\log\log\mu(\mathcal{F}))/\log\mu(\mathcal{F}))$ and $\mathrm{O}((\log\log\rho(\mathcal{F}))/\log\rho(\mathcal{F}))$ on robustness. Our randomized strategies with robustness $\Omega(1/\log\mu(\mathcal{F}))$ and $\Omega(1/\log\rho(\mathcal{F}))$ appear in § 3. In § 4, we discuss general independence systems. Section 5 concludes this paper with a few remarks.

2. Upper Bounds on Robustness

As we described in § 1.2, there exists a randomized strategy with robustness at least $1/\ln 4$ for bit-concave independence systems [11]. In this section, we show that there exists an instance of the knapsack problem for which no randomized strategy can achieve a constant robustness.

Theorem 1. *For an arbitrary constant $\kappa > 0$, there exists an instance of the knapsack problem such that the robustness of an arbitrary randomized strategy is less than κ.*

Proof. For a given constant $\kappa > 0$, let M and T be integers larger than $3/\kappa$. Consider the following instance of the knapsack problem (see Table 1).

- There are $T+1$ types of items, say type 0, type 1, ..., type T.
- For each $i = 0, 1, \ldots, T$, there are M^{2i} items of type i, and the weight and profit of each item of type i are M^{2T-2i} and M^{2T-i}, respectively.
- The capacity is $C = M^{2T}$.

Observe that the total weight of the items of type i is equal to C for each i. Since the density p_e/w_e of an item e of type i becomes larger for large i, it is better to choose items of type i with large i under a large cardinality bound. However, the profit of a single item of type i is small for large i, and hence it is better to choose items with small i under a small cardinality bound. For this instance, we show that the robustness of an arbitrary randomized strategy is less than κ.

Let $\Delta \subseteq \mathbf{R}_+^{T+1}$ be the set of all vectors $\delta = (\delta_0, \delta_1, \dots, \delta_T) \in \mathbf{R}_+^{T+1}$ such that $\delta_i M^{2i}$ is an integer for $i = 0, 1, \dots, T$ and $\sum_i \delta_i \leq 1$. For $\delta \in \Delta$, let $X_\delta \subseteq E$ denote the feasible solution of the knapsack instance that contains $\delta_i M^{2i}$ items of type i for $i = 0, 1, \dots, T$. Note that $\sum_i \delta_i \leq 1$ corresponds to the capacity constraint and there is a one-to-one correspondence between Δ and the set of all feasible solutions.

Since the set of all items of type i is a feasible solution, we have that $p(\mathrm{OPT}_{M^{2i}}) \geq M^{2T+i}$ for each $i = 0, 1, \dots, T$. For each $\delta \in \Delta$ and for each $i \in \{0, 1, \dots, T\}$, it holds that

$$p(X_\delta(M^{2i})) \leq \sum_{j=0}^{i-1} \delta_j M^{2j} \cdot M^{2T-j} + \delta_i M^{2i} \cdot M^{2T-i} + M^{2i} \cdot M^{2T-i-1},$$

where the last term bounds the total profit of the items of types $i+1, i+2, \dots, T$ in $X_\delta(M^{2i})$, because each profit is at most M^{2T-i-1} and the number of items is at most M^{2i}. The right-hand side of this inequality is bounded by

$$\left(\sum_{j=0}^{i-1} \delta_j\right) M^{2T+i-1} + \delta_i M^{2T+i} + M^{2T+i-1} \leq \delta_i M^{2T+i} + 2M^{2T+i-1},$$

which shows that

$$p(X_\delta(M^{2i})) \leq \left(\delta_i + \frac{2}{M}\right) \cdot p(\mathrm{OPT}_{M^{2i}}) \quad (i = 0, 1, \dots, T).$$

Hence, for a randomized strategy choosing X_δ with probability λ_δ, it holds that

$$\sum_{\delta \in \Delta} \lambda_\delta p(X_\delta(M^{2i})) \leq \left(\sum_{\delta \in \Delta} \lambda_\delta \delta_i + \frac{2}{M}\right) \cdot p(\mathrm{OPT}_{M^{2i}}) \quad (i = 0, 1, \dots, T),$$

which implies that the robustness of this strategy is at most $\min_i\{\sum_{\delta \in \Delta} \lambda_\delta \delta_i + (2/M)\}$. Since

$$\sum_{i=0}^{T} \left(\sum_{\delta \in \Delta} \lambda_\delta \delta_i\right) = \sum_{\delta \in \Delta} \left(\lambda_\delta \sum_{i=0}^{T} \delta_i\right) \leq \sum_{\delta \in \Delta} \lambda_\delta = 1,$$

it follows that $\min_i\{\sum_{\delta \in \Delta} \lambda_\delta \delta_i\} \leq 1/(T+1)$. Therefore, the robustness is at most $\frac{1}{T+1} + \frac{2}{M}$, which is smaller than κ. □

Since Theorem 1 shows that no randomized strategy can achieve a constant robustness, a reasonable objective is to obtain a good robustness in terms of some parameters, and we use the parameters $\mu(\mathcal{F})$ and $\rho(\mathcal{F})$, defined in (1) and (2), respectively. The proof of Theorem 1 shows that, for the instance in Table 1, the robustness of an arbitrary randomized strategy is both $\mathrm{O}((\log\log \rho(\mathcal{F}))/\log \rho(\mathcal{F}))$ and $\mathrm{O}((\log\log \mu(\mathcal{F}))/\log \mu(\mathcal{F}))$.

Theorem 2. *There exists a sequence of independence systems* $(E, \mathcal{F})$ *defined by instances of the knapsack problem such that the robustness of an arbitrary randomized strategy is both* $\mathrm{O}((\log\log \mu(\mathcal{F}))/\log \mu(\mathcal{F}))$ *and* $\mathrm{O}((\log\log \rho(\mathcal{F}))/\log \rho(\mathcal{F}))$.

Proof. Let $T = M$ in Table 1. Then, $\mu(\mathcal{F}) = \rho(\mathcal{F}) = M^{2M}$ and the robustness of an arbitrary randomized strategy is at most $3/M$. Since $\log M^{2M} = \Theta(M \log M)$ and $\log\log M^{2M} = \Theta(\log M)$, the theorem follows. $\square$

Parameters $\mu(\mathcal{F})$ and $\rho(\mathcal{F})$ represent the intractability of the independence system $(E, \mathcal{F})$ in a sense that they are closely related to the approximation ratio of the greedy algorithm. To see this, for an arbitrarily large integer L, consider an instance of the knapsack problem in which $E = \{e_0, e_1, \ldots, e_L\}$, $w_{e_0} = C$, $w_{e_i} = C/L$ for $i = 1, \ldots, L$, and $p_{e_i} = 1$ for $i = 0, 1, \ldots, L$. In this instance, $\mu(\mathcal{F}) = \rho(\mathcal{F}) = L$ and the greedy algorithm may return $\{e_0\}$, whose approximation ratio is no better than $1/L = 1/\mu(F) = 1/\rho(\mathcal{F})$.

On the other hand, Mestre [12] proved that the approximation ratio $1/\mu(\mathcal{F})$ of the greedy algorithm is guaranteed. We prove that the greedy algorithm attains $(1/\rho(\mathcal{F}))$-approximation as well. Note that $\rho(\mathcal{F})$ is a newly introduced parameter different from the exchangeability $\mu(\mathcal{F})$ and the rank quotient $r(\mathcal{F})$.

Proposition 3. *Let* $(E, \mathcal{F})$ *be an independence system and* $p \in \mathbf{R}_+^E$ *be a profit vector. For the problem of finding a feasible set* $X \in \mathcal{F}$ *maximizing* $p(X)$*, the greedy algorithm finds a* $(1/\rho(\mathcal{F}))$*-approximate solution.*

Proof. Let Y and OPT be the output of the greedy algorithm and an optimal solution, respectively. By the definition of $a_{\min}$, Y contains $a_{\min}$ highest profit elements in E, that is, $E(a_{\min}) \subseteq Y$. Recall that $X(k)$ is a subset of X satisfying that $|X(k)| \le k$ and $p(X(k)) = \max\{p(X') \mid X' \subseteq X, |X'| \le k\}$. Let $p_0 := \min\{p_e \mid e \in E(a_{\min})\}$. Since $p_{e'} \le p_0$ for each $e' \in \mathrm{OPT} - \mathrm{OPT}(a_{\min})$ and $|\mathrm{OPT}| \le a_{\max}$, we have

$$\begin{aligned} p(\mathrm{OPT}) &= p(\mathrm{OPT}(a_{\min})) + p(\mathrm{OPT} - \mathrm{OPT}(a_{\min})) \\ &\le p(E(a_{\min})) + (|\mathrm{OPT}| - a_{\min})p_0 \\ &\le p(E(a_{\min})) + (a_{\max} - a_{\min}) \cdot \frac{p(E(a_{\min}))}{a_{\min}} \\ &= \rho(\mathcal{F}) \cdot p(E(a_{\min})) \\ &\le \rho(\mathcal{F}) \cdot p(Y), \end{aligned}$$

showing that Y is a $(1/\rho(\mathcal{F}))$-approximate solution. $\square$

We close this section with remarking that the ratio $\mu(\mathcal{F})/\rho(\mathcal{F})$ can be arbitrarily large and small. To see this, consider an instance of the knapsack problem in which $C = 2M$, there is one item of weight M, and there are $2M$ items of weight 1. In this instance, $\mu(\mathcal{F}) = M$ and $\rho(\mathcal{F}) = 2M/(M+1) < 2$, which shows that $\mu(\mathcal{F})/\rho(\mathcal{F})$ can be arbitrarily large. Also, consider an instance in which $C = 2M - 1$, there are two items of weight M, and there are M items of weight 1. In this instance, $\mu(\mathcal{F}) = 1$ and $\rho(\mathcal{F}) = M$, showing that $\mu(\mathcal{F})/\rho(\mathcal{F})$ can be arbitrarily small.

3. Randomized Strategies

In Theorem 2, we have seen that there exist instances such that the optimal randomized strategy for the robust knapsack problem can be as small as $\mathrm{O}((\log\log\mu(\mathcal{F}))/\log\mu(\mathcal{F}))$ and $\mathrm{O}((\log\log\rho(\mathcal{F}))/\log\rho(\mathcal{F}))$. This section is devoted to presenting positive results, randomized strategies with robustness $\Omega(1/\log\rho(\mathcal{F}))$ and $\Omega(1/\log\mu(\mathcal{F}))$ in § 3.1 and 3.2, respectively. Theorem 2 suggests that these results are almost tight, and the latter robustness substantially improves upon the robustness $1/\sqrt{\mu(\mathcal{F})}$ of a deterministic strategy in [6].

3.1. An $\Omega(1/\log\rho(\mathcal{F}))$-robust strategy

In this subsection, we present a randomized strategy with robustness $\Omega(1/\log\rho(\mathcal{F}))$, which can be applied to general independence systems. Recall that $\rho(\mathcal{F})$ is defined in (2).

Theorem 4. *For an arbitrary independence system $(E,\mathcal{F})$ with a profit vector $p\in\mathbf{R}_+^E$, there is a randomized strategy with robustness $\Omega(1/\log\rho(\mathcal{F}))$.*

Proof. Let $m=\lceil\log\rho(\mathcal{F})\rceil$. Recall that, for each k, OPT_k is a maximizer of $p(X)$ subject to $X\in\mathcal{F}$ and $|X|\le k$. Our randomized strategy is described as follows.

Strategy 1. Choose $X_i := \mathrm{OPT}_{2^i a_{\min}}$ with probability $1/(m+1)$ for each $i\in\{0,1,\dots,m\}$.

- For an integer k with $a_{\min}\le k<2^m a_{\min}$, let j be the unique integer satisfying $2^j a_{\min}\le k<2^{j+1}a_{\min}$. Then, we have that

$$\begin{aligned}p(X_j(k)) &= p\left(\mathrm{OPT}_{2^j a_{\min}}\right)\\ &\ge p(\mathrm{OPT}_k(2^j a_{\min}))\\ &\ge \frac{2^j a_{\min}}{k}\cdot p(\mathrm{OPT}_k)\\ &\ge \frac{1}{2}\cdot p(\mathrm{OPT}_k).\end{aligned}$$

We also have that

$$p(X_{j+1}(k))\ge\frac{k}{2^{j+1}a_{\min}}\cdot p(X_{j+1})\ge\frac{k}{2^{j+1}a_{\min}}\cdot p(\mathrm{OPT}_k)\ge\frac{1}{2}\cdot p(\mathrm{OPT}_k).$$

Thus,

$$\begin{aligned}\mathbf{E}[p(X(k))] &= \frac{1}{m+1}\sum_{i=0}^{m}p(X_i(k))\\ &\ge\frac{1}{m+1}\cdot(p(X_j)+p(X_{j+1}(k)))\\ &\ge\frac{1}{m+1}\cdot p(\mathrm{OPT}_k).\end{aligned}$$

- For an integer $k \leq a_{\min}$, we have $p(X_0(k)) = p(\mathrm{OPT}_k)$, since $X_0 = \mathrm{OPT}_{a_{\min}}$ is the set of $a_{\min}$ highest profit elements in E. Thus,

$$\mathbf{E}[p(X(k))] \geq \frac{1}{m+1} \cdot p(X_0(k)) = \frac{1}{m+1} \cdot p(\mathrm{OPT}_k).$$

- For an integer $k \geq 2^m a_{\min}$, we have that $2^m a_{\min} \geq a_{\max}$ by the definition of m, and hence $p(\mathrm{OPT}_k) = p(X_m) = p(X_m(k))$. Thus,

$$\mathbf{E}[p(X(k))] \geq \frac{1}{m+1} \cdot p(X_m(k)) = \frac{1}{m+1} \cdot p(\mathrm{OPT}_k).$$

Therefore, we conclude that the robustness of Strategy 1 is at least $1/(m+1)$. □

In order to obtain the strategy in polynomial time, we have to compute $a_{\min}$ and $\mathrm{OPT}_{2^i a_{\min}}$ efficiently. If $(E, \mathcal{F})$ is defined by an instance of the knapsack problem, we can easily compute $a_{\min}$ and $a_{\max}$ in polynomial time. Although computing $\mathrm{OPT}_{2^i a_{\min}}$ is NP-hard, we can efficiently compute a solution X_i approximating $\mathrm{OPT}_{2^i a_{\min}}$ for each i via an FPTAS for the knapsack problem with a cardinality constraint [1]. Hence, we obtain the following corollary.

Corollary 5. *For an arbitrary independence system $(E, \mathcal{F})$ defined by an instance of the knapsack problem, an $\Omega(1/\log \rho(\mathcal{F}))$-robust randomized strategy is obtained in polynomial time.*

We also note that we can slightly improve the bound by slightly modifying the proof of Theorem 4. Let $a^*_{\max}$ be the minimum size of an optimal solution of the knapsack problem. Then, we can replace $a_{\max}$ with $a^*_{\max}$ in the proof to obtain an $\Omega(1/\log(a^*_{\max}/a_{\min}))$-robust strategy, which is slightly better than $\Omega(1/\log \rho(\mathcal{F}))$.

3.2. An $\Omega(1/\log \mu(\mathcal{F}))$-robust strategy

In this subsection, we present an $\Omega(1/\log \mu(\mathcal{F}))$-robust randomized strategy, where $\mu(\mathcal{F})$ is the exchangeability of the independence system $(E, \mathcal{F})$. Note that, for the case where only deterministic strategies are allowed, Kakimura and Makino [6] showed the existence of $(1/\sqrt{\mu(\mathcal{F})})$-robust solution. That is, we improve this ratio to $\Omega(1/\log \mu(\mathcal{F}))$ by allowing randomized strategies, to prove the power of randomness in the robust knapsack problem. Our strategy is based on the ideas in § 3.1, but we need extra work for this case.

Theorem 6. *For an arbitrary independence system $(E, \mathcal{F})$ defined by an instance of the knapsack problem, there is a randomized strategy with robustness $\Omega(1/\log \mu(\mathcal{F}))$.*

Proof. Let $(E, \mathcal{F})$ be defined by an instance (E, p, w, C) of the knapsack problem. In this proof we often abbreviate $\mu(\mathcal{F})$ as μ. Recall that we assume $w_e \leq C$ for each $e \in E$ and $w(E) > C$. Let $Y \subseteq E$ be an optimal solution of this problem, and let $Z \subseteq E$ be the set of $a_{\min}$ heaviest elements in E. Note that $w(Z) \leq C$ by the definition of $a_{\min}$.

We cannot apply Strategy 1 directly, because $|Y|$ might be much larger than $\mu|Z|$. To overcome this difficulty, we choose many light elements in Y in advance (ignoring their profit), which is our main idea in the proof. Let Y_0 be the subset of Y that maximizes $|Y_0|$ subject to $w(Y_0) \leq C - w(Z)$. That is, Y_0 is obtained by taking light elements in Y greedily as long as $w(Y_0) \leq C - w(Z)$. Now we have the following lemma.

Lemma 7. *It holds that* $\mu|Z| \geq |Y - Y_0|$.

Proof of Lemma 7. We first show the existence of a feasible set $Y^* \subseteq Y \cup Z$ such that $Z \subseteq Y^*$ and $|Y^* - Z| \geq |Y - Z| - \mu|Z - Y|$. If $Z - Y = \emptyset$, then $Y^* = Y$ satisfies these conditions. Otherwise, let z be an element in $Z - Y$, and apply (1) between $Y, Z \in \mathcal{F}$ with respect to $z \in Z - Y$. Then, by the definition of μ, there exists a feasible set $Y' \subseteq Y \cup \{z\}$ such that $(Y \cap Z) \cup \{z\} \subseteq Y'$ and $|Y - Y'| \leq \mu$. That is, if we replace Y with Y', then $|Z - Y|$ decreases by one and $|Y - Z|$ decreases at most μ. Next, we apply the exchange between Y' and Z to obtain Y''. By repeating this procedure $|Z - Y|$ times, we obtain a feasible set $Y^* \subseteq Y \cup Z$ such that $Z \subseteq Y^*$ and

$$|Y^* - Z| \geq |Y - Z| - \mu|Z - Y|. \tag{3}$$

Since $Z \subseteq Y^*$ implies that $w(Y^* - Z) \leq C - w(Z)$, it holds that $|Y_0| \geq |Y^* - Z|$ by the definition of Y_0. By combining this with (3), we have $|Y_0| \geq |Y - Z| - \mu|Z - Y|$, which is equivalent to $\mu|Z - Y| \geq |Y - Z| - |Y_0|$. By adding $\mu|Y \cap Z| \geq |Y \cap Z|$ to this inequality, we obtain $\mu|Z| \geq |Y - Y_0|$. (End of the proof of Lemma 7)

Define $C' := C - w(Y_0)$, $E' := E - Y_0$, and $m' := \lceil \log(|Y - Y_0|/a_{\min}) \rceil$. Then, $m' = \mathrm{O}(\log \mu)$ by Lemma 7 and $a_{\min} = |Z|$. Consider the instance (E', p, w, C') of the knapsack problem, where p and w are restricted to E'. For each k, let OPT'_k be an optimal solution of (E', p, w, C') subject to $|\mathrm{OPT}'_k| \leq k$.

The following lemma plays an important role in our algorithm.

Lemma 8. *For an arbitrary $X \subseteq E$ with $w(X) \leq C$, X can be partitioned into three sets X^1, X^2, and X^3 so that $w(X^\ell) \leq C'$ for $\ell = 1, 2, 3$ (possibly $X^\ell = \emptyset$), where C' is defined as above.*

Proof of Lemma 8. We first observe that $C' = C - w(Y_0) \geq w(Z) \geq C/2$ and there is no element in X whose weight is greater than C'.

If $w(X) \leq C'$, then the lemma is obvious. Otherwise, define X^1, X^2, and X^3 as follows.

- Let X^1 be a maximal subset of X satisfying that $w(X^1) \leq C'$.
- Let $X^2 = \{x\}$ for some $x \in X - X^1$.
- Let $X^3 = X - (X^1 \cup X^2)$.

Then, it is clear that $w(X^1) \le C'$ and $w(X^2) \le C'$. Furthermore, since $w(X^1 \cup X^2) > C'$ by the maximality of X^1, it follows that $w(X^3) = w(X) - w(X^1 \cup X^2) < w(X) - C' \le C'$ from $C' \ge C/2$. (End of the proof of Lemma 8)

Our randomized strategy is described as follows.

Strategy 2. Choose $X_i := \mathrm{OPT}'_{2^i a_{\min}} \cup Y_0$ with probability $1/(m'+1)$ for each $i \in \{0, 1, \ldots, m'\}$.

We now analyze the robustness of Strategy 2. To simplify the notation, let $X'_i := \mathrm{OPT}'_{2^i a_{\min}}$ for each i.

- For an integer k with $a_{\min} \le k < 2^{m'} a_{\min}$, let j be the unique integer satisfying $2^j a_{\min} \le k < 2^{j+1} a_{\min}$. Then, it holds that

$$\begin{aligned} p(X_{j+1}(k)) &\ge p(X'_{j+1}(k)) \\ &\ge \frac{k}{2^{j+1} a_{\min}} \cdot p(X'_{j+1}) \\ &\ge \frac{k}{2^{j+1} a_{\min}} \cdot p(\mathrm{OPT}'_k) \\ &\ge \frac{1}{2} \cdot p(\mathrm{OPT}'_k). \end{aligned} \tag{4}$$

  By Lemma 8, $\mathrm{OPT}_k - Y_0$ can be partitioned into three sets OPT^1_k, OPT^2_k, and OPT^3_k so that $w(\mathrm{OPT}^\ell_k) \le C'$ for $\ell = 1, 2, 3$, which shows that

$$\begin{aligned} p(\mathrm{OPT}_k) &= p(\mathrm{OPT}_k - Y_0) + p(\mathrm{OPT}_k \cap Y_0) \\ &\le p(\mathrm{OPT}^1_k) + p(\mathrm{OPT}^2_k) + p(\mathrm{OPT}^3_k) + p(Y_0(k)) \\ &\le 3p(\mathrm{OPT}'_k) + p(X_{j+1}(k)). \end{aligned} \tag{5}$$

  By (4) and (5), we have that $p(\mathrm{OPT}_k) \le 7p(X_{j+1}(k))$. Thus,

$$\begin{aligned} \mathbf{E}[p(X(k))] &= \frac{1}{m'+1} \sum_{i=0}^{m'} p(X_i(k)) \\ &\ge \frac{1}{m'+1} \cdot p(X_{j+1}(k)) \\ &\ge \frac{1}{7(m'+1)} \cdot p(\mathrm{OPT}_k). \end{aligned}$$

- For an integer $k \le a_{\min}$, we have $p(X_0(k)) = p(\mathrm{OPT}_k)$, since X'_0 is the set of $a_{\min}$ highest profit elements in $E' = E - Y_0$. Thus,

$$\mathbf{E}[p(X(k))] \ge \frac{1}{m'+1} \cdot p(X_0(k)) = \frac{1}{m'+1} \cdot p(\mathrm{OPT}_k).$$

- For an integer $k \ge 2^{m'} a_{\min}$, we note that $p(\mathrm{OPT}'_k) = p(Y - Y_0) = p(X'_{m'}) = p(X'_{m'}(k))$. By Lemma 8, $\mathrm{OPT}_k - Y_0$ can be partitioned into three sets OPT_k^1, OPT_k^2, and OPT_k^3 so that $w(\mathrm{OPT}_k^\ell) \le C'$ for $\ell = 1, 2, 3$, which shows that

$$\begin{aligned} p(\mathrm{OPT}_k) &= p(\mathrm{OPT}_k - Y_0) + p(\mathrm{OPT}_k \cap Y_0) \\ &\le p(\mathrm{OPT}_k^1) + p(\mathrm{OPT}_k^2) + p(\mathrm{OPT}_k^3) + p(Y_0(k)) \\ &\le 3p(\mathrm{OPT}'_k) + p(X_{m'}(k)) \\ &= 4p(X_{m'}(k)). \end{aligned}$$

Thus,

$$\mathbf{E}[p(X(k))] \ge \frac{1}{m'+1} \cdot p(X_{m'}(k)) = \frac{1}{4(m'+1)} \cdot p(\mathrm{OPT}_k).$$

Therefore, we conclude that the robustness of Strategy 2 is at least $1/7(m'+1) = \Omega(1/\log\mu)$. □

In the same way as Corollary 5, since we can efficiently compute a solution X_i approximating $\mathrm{OPT}'_{2^i a_{\min}}$ for each i via an FPTAS [1], we obtain the following corollary to Theorem 6.

Corollary 9. *For an arbitrary independence system $(E, \mathcal{F})$ defined by an instance of the knapsack problem, an $\Omega(1/\log\mu(\mathcal{F}))$-robust randomized strategy is obtained in polynomial time.*

4. Extension to General Independence Systems

In this section, we extend our results to general independence systems. We have already seen in Theorem 4 that Strategy 1 can be applied to general independence systems and its robustness is $\Omega(1/\log\rho(\mathcal{F}))$. Furthermore, if $a_{\min}, a_{\max}$, and OPT_k are (approximately) computable in polynomial time, then we can obtain the strategy in polynomial time.

In what follows, we show robustness in general independence systems. More precisely, we improve the upper bounds given in Theorem 2 to $\mathrm{O}(1/\log\mu(\mathcal{F}))$ and $\mathrm{O}(1/\log\rho(\mathcal{F}))$ for general independence systems.

Theorem 10. *There exists a sequence of independence systems $(E, \mathcal{F})$ such that the robustness of an arbitrary randomized strategy is both $\mathrm{O}(1/\log\mu(\mathcal{F}))$ and $\mathrm{O}(1/\log\rho(\mathcal{F}))$.*

Proof. Let M be a constant larger than 1 (e.g., $M = 10$), and consider the following independence system $(E, \mathcal{F})$ (see Table 2).

- The set E consists of $T+1$ types of items, say type 0, type 1, ..., type T.
- For each $i = 0, 1, \ldots, T$, type i has M^{2i} items with profit M^{2T-i}.
- $\mathcal{F}$ is the collection of all the subsets of E consisting of at most one type of items.

Table 2: An independence system with small robustness. A set of items is feasible if it consists of at most one type of items.

type	p	number of items	total profit
0	M^{2T}	1	M^{2T}
1	M^{2T-1}	M^2	M^{2T+1}
2	M^{2T-2}	M^4	M^{2T+2}
	$\vdots$		
i	M^{2T-i}	M^{2i}	M^{2T+i}
	$\vdots$		
$T-1$	M^{T+1}	M^{2T-2}	M^{3T-1}
T	M^{T}	M^{2T}	M^{3T}

It is not difficult to see that $\rho(\mathcal{F}) = \mu(\mathcal{F}) = M^{2T}$ for this independence system. We show that the robustness of an arbitrary randomized strategy is $\mathrm{O}(1/T)$.

For $i = 0, 1, \ldots, T$, let X_i be the feasible set consisting of all M^{2i} items of type i. By the definition of $\mathcal{F}$, $\{X_0, X_1, \ldots, X_T\}$ is the set of all maximal feasible sets, and hence it suffices to consider a randomized strategy choosing $X_0, X_1, \ldots, X_T$.

For $i, j \in \{0, 1, \ldots, T\}$, we have that $p(X_j(M^{2i})) = M^{2T+i-|i-j|}$. Consider a randomized strategy choosing X_j with probability λ_j. Since $p(\mathrm{OPT}_{M^{2i}}) = p(X_i) = M^{2T+i}$, it follows that

$$\sum_{j=0}^{T} \lambda_j p(X_j(M^{2i})) = \left(\sum_{j=0}^{T} \lambda_j M^{-|i-j|}\right) \cdot p(\mathrm{OPT}_{M^{2i}}) \quad (i = 0, 1, \ldots, T),$$

which implies that the robustness of this strategy is at most $\min_i \left\{\sum_{j=0}^{T} \lambda_j M^{-|i-j|}\right\}$. Since

$$\begin{aligned} \sum_{i=0}^{T}\left(\sum_{j=0}^{T} \lambda_j M^{-|i-j|}\right) &= \sum_{j=0}^{T} \lambda_j \left(\sum_{i=0}^{T} M^{-|i-j|}\right) \\ &\leq \sum_{j=0}^{T} \lambda_j \left(1 + 2\sum_{i'=1}^{\infty} M^{-i'}\right) \\ &\leq 1 + \frac{2}{M-1} = \mathrm{O}(1), \end{aligned}$$

the robustness is at most $\min_i \left\{\sum_{j=0}^{T} \lambda_j M^{-|i-j|}\right\} = \mathrm{O}(1/T)$, which completes the proof. □

Theorem 10 shows that the robustness $\Omega(1/\log \rho(\mathcal{F}))$ given in Theorem 4 is tight when we consider general independence systems.

5. Concluding Remarks

In this paper, we have addressed randomized strategies for the robust independence systems defined by the knapsack problem. We exhibited upper bounds on robustness in terms of the exchangeability $\mu(\mathcal{F})$ and a newly introduced parameter $\rho(\mathcal{F})$, which represent the intractability of the independence system $(E, \mathcal{F})$. We then designed randomized strategies with better robustness than deterministic strategies, and extended those results to general independence systems.

A major task for future research would be filling the gap between the upper and lower bounds on robustness. Extending Theorem 6, a lower bound in terms of the exchangeability $\mu(\mathcal{F})$, to general independence systems, and providing upper or lower bounds in terms of the rank quotient $r(\mathcal{F})$ are also of interest.

References

[1] A. Caprara, H. Kellerer, U. Pferschy, and D. Pisinger, *Approximation algorithms for knapsack problems with cardinality constraints*, European Journal of Operations Research, 123 (2000), pp. 333–345.

[2] Y. Disser, M. Klimm, N. Megow, and S. Stiller, *Packing a knapsack of unknown capacity*, in Proceedings of the 31st International Symposium on Theoretical Aspects of Computer Science (STACS 2014), 2014, pp. 276–287.

[3] R. Fujita, Y. Kobayashi, and K. Makino, *Robust matchings and matroid intersections*, SIAM Journal on Discrete Mathematics, 27 (2013), pp. 1234–1256.

[4] R. Hassin and S. Rubinstein, *Robust matchings*, SIAM Journal on Discrete Mathematics, 15 (2002), pp 530–537.

[5] T.A. Jenkyns, *The efficiency of the greedy algorithm*, in Proceedings of the 7th Southeastern Conference on Combinatorics, Graph Theory, and Computing, Unitas Mathematica, 1976, pp. 341–350.

[6] N. Kakimura and K. Makino, *Robust independence systems*, SIAM Journal on Discrete Mathematics, 27 (2013), pp. 1257–1273.

[7] N. Kakimura, K. Makino, and K. Seimi, *Computing knapsack solutions with cardinality robustness*, Japan Journal of Industrial and Applied Mathematics, 29 (2012), pp. 469–483.

[8] Y. Kobayashi and K. Takazawa, *Randomized strategies for cardinality robustness in the knapsack problem*, in Proceedings of the 13th Meeting on Analytic Algorithmics and Combinatorics (ANALCO 2016), 2016, pp. 25–33.

[9] B. Korte and D. Hausmann, *An analysis of the greedy heuristic for independence systems*, Annals of Discrete Mathematics, 2 (1978), pp. 65–74.

[10] A. Mastin, P. Jaillet, and S. Chin, *Randomized minmax regret for combinatorial optimization under uncertainty*, in Proceedings of the 26th International Symposium on Algorithms and Computation (ISAAC 2015), 2015, pp. 491–501.

[11] J. Matuschke, M. Skutella, and J.A. Soto, *Robust randomized matchings*, in P. Indyk, ed., in Proceedings of the 26th Annual ACM-SIAM Symposium on Discrete Algorithms, 2015, pp. 1904–1915.

[12] J. Mestre, *Greedy in approximation algorithms*, in Y. Azar and T. Erlebach, eds., in Proceedings of the 14th Annual European Symposium on Algorithms, LNCS 4168, Springer-Verlag, 2006, pp. 528–539.