Randomized Strategies for Cardinality Robustness in the Knapsack Problem

Yusuke Kobayashi[*] Kenjiro Takazawa[†]

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 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 solution under a cardinality constraint is computable.

1 Introduction

1.1 Cardinality robustness in independence systems *Cardinality robustness* in independence systems is introduced by Hassin and Rubinstein [3], 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 $\mathrm{OPT}_k \in \mathcal{F}$, $|\mathrm{OPT}_k| \le k$, and $p(\mathrm{OPT}_k) = \max\{p(X) \mid X \in \mathcal{F}, |X| \le k\}$, 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 $|X(k)| \le k$ and $p(X(k)) = \max\{p(X') \mid X' \subseteq X, |X'| \le k\}$. 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 $p(X(k)) \ge \alpha \cdot p(\mathrm{OPT}_k)$ 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[1] is 1-robust. More generally, Hassin and Rubinstein [3] proved that a greedy solution is $r(\mathcal{F})$-robust, where $r(\mathcal{F})$ is the *rank quotient* of $(E, \mathcal{F})$ [4, 7].

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 [3] 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 [2] 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 [2] that determining whether a graph has an α-robust matching is NP-hard for an arbitrary $\alpha > 1/\sqrt{2}$. Analysis for general independence systems is due to Kakimura and Makino [5], 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

[*]University of Tsukuba. Supported by JST, ERATO, Kawarabayashi Project, by JSPS KAKENHI Grant Number 24700004, and by MEXT KAKENHI Grant Number 24106002. E-mail: `kobayashi@sk.tsukuba.ac.jp`

[†]Research Institute for Mathematical Sciences, Kyoto University. Supported by JSPS KAKENHI Grant Numbers 25280004 and 26280001. E-mail: `takazawa@kurims.kyoto-u.ac.jp`

[1]The *greedy algorithm* for an independence system is defined as follows. Sort the elements $e \in E$ by profit p_e, i.e., $E = \{e_1, \dots, e_n\}$ and $p_{e_1} \ge \cdots \ge p_{e_n}$. Let $X_0 = \emptyset$, and for $i = 1, \dots, n$, 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.*

integer μ satisfying that

$$\forall X, Y \in \mathcal{F},\ \forall e \in Y - X, \\ \exists Z \subseteq X - Y \text{ s.t. } |Z| \le \mu,\ (X - Z) \cup \{e\} \in \mathcal{F}. \tag{1.1}$$

In [5], 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 [6] 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., $\mathcal{F} = \{X \subseteq E \mid w(X) \le C\}$. Kakimura, Makino, and Seimi [6] 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.

1.2 Randomized strategies

The above results correspond to deterministic strategies (or pure strategies) of the zero-sum game. Matuschke, Skutella, and Soto [8] 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 [8] 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, which are defined as follows.

> If $p_e = 2^{l_e}$ for each $e \in E$ with $l_e \in \mathbf{Z}$ (i.e., p is a *bit-function*), then a greedy solution is 1-robust. Equivalently, for an arbitrary bit-function p, it holds that $2p(\mathrm{OPT}_{k+1}) \ge p(\mathrm{OPT}_k) + p(\mathrm{OPT}_{k+2})$ for all positive integer k (*bit-concavity*).

Examples of bit-concave independence systems include matroid intersection, b-matchings, strongly base orderable matchoids, strongly base orderable matroid parity systems.

1.3 Our results

We address randomized strategies for the robust independence systems defined by an instance of the knapsack problem. It is not difficult to see that those independence systems are not necessarily bit-concave, and hence the method in [8] cannot applied. In what follows, we assume that $\mathcal{F} \ne 2^E$ and $\{e\} \in \mathcal{F}$ for every $e \in E$. That is, $w_e \le C$ 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

$$\begin{aligned} \rho(\mathcal{F}) &:= \frac{a_{\max}}{a_{\min}}, \\ a_{\max} &:= \max\{|X| \mid X \in \mathcal{F}\}, \\ a_{\min} &:= \min\{|X| \mid X \notin \mathcal{F}\} - 1. \end{aligned} \tag{1.2}$$

We remark that the parameters $\mu(\mathcal{F})$ and $\rho(\mathcal{F})$ represent the intractability of the independence system $(E, \mathcal{F})$. Clearly $\mu(\mathcal{F}) \ge 1$ and $\rho(\mathcal{F}) \ge 1$, $\mu(\mathcal{F}) = 1$ 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 $\mu(\mathcal{F}) \le 2$. For the problem of finding a feasible set X maximizing $p(X)$, the greedy algorithm attains $1/\mu(\mathcal{F})$-approximation [9]. Further, we show that a greedy solution yields $1/\rho(\mathcal{F})$-approximation as well (see Proposition 2.1), and construct examples for which the greedy algorithm has approximation ratio no better than $1/\mu(\mathcal{F})$ and $1/\rho(\mathcal{F})$ (see Section 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 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. We show that the $\Omega(1/\log\rho(\mathcal{F}))$-robust strategy is extended to general independence systems. We also provide upper bounds $\mathrm{O}(1/\log\rho(\mathcal{F}))$ and $\mathrm{O}(1/\log\mu(\mathcal{F}))$ on robustness, which proves the tightness of our $\Omega(1/\log\rho(\mathcal{F}))$-robust strategy.

We also point out that an independence system defined by an instance (E, p, w, C) of the knapsack problem is an example of an independence system which is bit-concave but not concave, when all items have unit densities, i.e., p_e/w_e is constant. This provides an answer to a question posed by [8].

1.4 Organization of the paper The rest of this paper is organized as follows. In Section 2, we show an instance of the knapsack problem for which no randomized strategy attains constant 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 Section 3. Bit-concavity in the unit density case is also discussed in this section. In Section 4, we discuss general independence systems. Section 5 concludes this paper with a few remarks.

2 Upper Bounds on Robustness

As we described in Section 1.2, there exists a randomized strategy with robustness at least $1/\ln(4)$ for bit-concave independence systems [8]. 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 2.1. *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 $C = M^{2T}$.

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

type	w	p	number of items
0	M^{2T}	M^{2T}	1
1	M^{2T-2}	M^{2T-1}	M^2
2	M^{2T-4}	M^{2T-2}	M^4
		$\vdots$	
i	M^{2T-2i}	M^{2T-i}	M^{2i}
		$\vdots$	
$T-1$	M^2	M^{T+1}	M^{2T-2}
T	1	M^T	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 soft cardinality constraint. 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 hard cardinality constraint. 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 $\delta = (\delta_0, \delta_1, \ldots, \delta_T) \in \mathbf{R}_+^{T+1}$ such that $\delta_i M^{2i}$ is an integer for $i = 0, 1, \ldots, T$ 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 $i = 0, 1, \ldots, T$. 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 $i = 0, 1, \ldots, T$. For each $\delta \in \Delta$ and for each $i \in \{0, 1, \ldots, T\}$, 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 $i+1, i+2, \ldots, T$ 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

$$p(X_\delta(M^{2i}))\leq\left(\delta_i+\frac{2}{M}\right)\cdot p(\mathrm{OPT}_{M^{2i}})$$

for $i=0,1,\ldots,T$.

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

$$\sum_{\delta\in\Delta}\lambda_\delta p(X_\delta(M^{2i}))\leq\left(\sum_{\delta\in\Delta}\lambda_\delta\delta_i+\frac{2}{M}\right)\cdot p(\mathrm{OPT}_{M^{2i}})$$

for $i=0,1,\ldots,T$, 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 2.1 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 (1.1) and (1.2), respectively. The proof of Theorem 2.1 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 2.2. *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. □

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 [9] proved that the approximation ratio $1/\mu(\mathcal{F})$ of the greedy algorithm is guaranteed [9]. We prove that the greedy algorithm attains $1/\rho(\mathcal{F})$-approximation as well.

Proposition 2.1. *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, $E(a_{\min})\subseteq Y$. Let $p_0:=\min\{p_e\mid e\in E(a_{\min})\}$. Since $p_{e'}\leq p_0$ for each $e'\in\mathrm{OPT}-\mathrm{OPT}(a_{\min})$ and $|\mathrm{OPT}|\leq a_{\max}$, we have

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

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

We close this section with remarking that 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 2.2, we have seen that the robustness of an arbitrary 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 Sections 3.1 and 3.2, respectively. Theorem 2.2 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 [5]. We also show in Section 3.3 that the randomized strategy in [8] yields robustness $1/\ln(4)$ for the case when all items have unit densities, i.e., p_e/w_e is constant.

3.1 An $\Omega(1/\log\rho(\mathcal{F}))$-robust strategy In this subsection, we present a randomized strategy with robustness $\Omega(1/\log\rho(\mathcal{F}))$. Recall that $\rho(\mathcal{F})$ is defined in (1.2).

THEOREM 3.1. *For an arbitrary independence system $(E,\mathcal{F})$ defined by an instance of the knapsack problem, there is a randomized strategy with robustness $\Omega(1/\log\rho(\mathcal{F}))$.*

Proof. Let $(E,\mathcal{F})$ be defined by an instance (E,p,w,C) of the knapsack problem and let $m = \lceil\log\rho(\mathcal{F})\rceil$. Recall that, for each k, OPT_k is an optimal solution of (E,p,w,C) subject to $|\mathrm{OPT}_k|\le k$. 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 $i\in\{0,1,\dots,m\}$.

We now show that the robustness of Strategy 1 is at least $1/(m+1) = \Omega(1/\log\rho(\mathcal{F}))$.

- 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

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

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 $k\le a_{\min}$, 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,

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

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

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

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

We remark that computing $\mathrm{OPT}_{2^i a_{\min}}$ is NP-hard. In order to obtain the strategy in polynomial time, we 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].

COROLLARY 3.1. *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 following the proof of Theorem 3.1. 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 [5] 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 Section 3.1, but we need extra work for this case.

THEOREM 3.2. *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 $a_{\min}$.

Since $|Y|/|Z| \geq a^*_{\max}/a_{\min}$, we can apply Strategy 1 when $|Y|/|Z| \leq \mu$ (see a remark after Corollary 3.1). We now address the case when $|Y|/|Z|$ is much larger. In such a case, 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 3.1. *It holds that $\mu|Z| \geq |Y - Y_0|$.*

Proof of Lemma 3.1. 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 (1.1) 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

$$|Y^* - Z| \geq |Y - Z| - \mu|Z - Y|. \tag{3.3}$$

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 (3.3), 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 $\mu|Z| \geq |Y - Y_0|$. □□

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 3.1 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 3.2. *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 $X^\ell = \emptyset$).*

Proof of Lemma 3.2. 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 $w(X^1) \leq C'$ and $w(X^2) \leq C'$. Furthermore, since $w(X^1 \cup X^2) > C'$ by the maximality of X^1, it follows that $w(X^3) = w(X) - w(X^1 \cup X^2) < w(X) - C' \leq C'$ from $C' \geq C/2$. □□

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

$$\begin{aligned} p(X_{j+1}(k)) &\geq p(X'_{j+1}(k)) \\ &\geq \frac{k}{2^{j+1} a_{\min}} \cdot p(X'_{j+1}) \\ &\geq \frac{k}{2^{j+1} a_{\min}} \cdot p(\mathrm{OPT}'_k) \\ &\geq \frac{1}{2} \cdot p(\mathrm{OPT}'_k). \end{aligned} \tag{3.4}$$

By Lemma 3.2, $\mathrm{OPT}_k - Y_0$ can be partitioned into three sets OPT^1_k, OPT^2_k, and OPT^3_k so that $w(\mathrm{OPT}^\ell_k) \leq C'$ for $\ell = 1, 2, 3$, which shows that

$$\begin{aligned} p(\mathrm{OPT}_k) &= p(\mathrm{OPT}_k - Y_0) + p(\mathrm{OPT}_k \cap Y_0) \\ &\leq \sum_{\ell=1}^{3} p(\mathrm{OPT}^\ell_k) + p(Y_0(k)) \\ &\leq 3p(\mathrm{OPT}'_k) + p(X_{j+1}(k)). \end{aligned} \tag{3.5}$$

By (3.4) and (3.5), we have that $p(\mathrm{OPT}_k) \leq 7p(X_{j+1}(k))$. Thus,

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

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

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

- For an integer $k \geq 2^{m'} a_{\min}$, we note that $p(\mathrm{OPT}'_k) = p(Y - Y_0) = p(X'_{m'}) = p(X'_{m'}(k))$. By Lemma 3.2, $\mathrm{OPT}_k - Y_0$ can be partitioned into three sets OPT^1_k, OPT^2_k, and OPT^3_k so that $w(\mathrm{OPT}^\ell_k) \leq C'$ for $\ell = 1, 2, 3$, which shows that

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

  Thus,

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

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

3.3 Unit density case In this subsection, we show that an instance of the knapsack problem (E, p, w, C) defines a bit-concave independence system (see Section 1.2 for definition) if all items have unit densities, i.e., p_e/w_e is constant, and thus the $1/\ln(4)$-robust strategy in [8] works for this case.

Proposition 3.1. *If an independence system is defined by an instance (E, p, w, C) of the knapsack problem in which p_e/w_e is constant, then it is bit-concave. This implies that there is a randomized strategy with robustness $1/\ln(4)$.*

Proof. Let p be a bit-function, i.e., $p_e = 2^{l_e}$ for each $e \in E$ with $l_e \in \mathbf{Z}$. Without loss of generality, assume that $p_e/w_e = 1$ for each $e \in E$. It suffices to show that a greedy solution X for this problem is 1-robust. To derive a contradiction, assume that X is not 1-robust, and let k be the minimum integer such that $p(X(k)) < p(\mathrm{OPT}_k)$. Let $Z := X(k) \cap \mathrm{OPT}_k$ and let e_0 be the cheapest element in $\mathrm{OPT}_k - Z$. Note that $X(k)$ and OPT_k are distinct solutions of the same size k, and hence $\mathrm{OPT}_k - Z \neq \emptyset$. We consider the following two cases separately.

- Consider the case when $p_e \geq p_{e_0}$ for every $e \in X(k) - Z$. Let $Z' := \{e' \in Z \mid p_{e'} \leq p_{e_0}\}$. Then

$$\begin{aligned}p(X(k - |Z'| - 1)) &\leq p(X(k) - Z') - p_{e_0} \\ &< p(\mathrm{OPT}_k - Z') - p_{e_0} \\ &\leq p(\mathrm{OPT}_{k-|Z'|-1}).\end{aligned} \tag{3.6}$$

  The first line follows from $X(k) - Z' - \{e\}$ for $e \in X(k) - Z'$ is a subset of X of size $k - |Z'| - 1$ and profit at most the right-hand side. The second line follows because $Z' \subseteq Z = X(k) \cap \mathrm{OPT}_k$ and $p(X(k)) < p(\mathrm{OPT}_k)$. Finally, the third line follows from $\mathrm{OPT}_k - Z' - \{e_0\}$ is a solution of size $k - |Z'| - 1$. Now, (3.6) contradicts the minimality of k.

- Consider the case when there exists $e' \in X(k) - Z$ with $p_{e'} < p_{e_0}$. Let $X' := \{e \in X(k) \mid p_e \geq p_{e_0}\}$. Then the existence of e' implies that $|X'| \leq k - 1$. By the definition of bit-functions, both $p(X' - Z)$ and $p(\mathrm{OPT}_k - Z)$ are integral multiples of p_{e_0}, which shows that $p(X' - Z) \leq p(\mathrm{OPT}_k - Z) - p_{e_0}$. Since $w = p$, we obtain $w(X') \leq w(\mathrm{OPT}_k) - w_{e_0} \leq C - w_{e_0}$, which contradicts that e_0 is not contained in the greedy solution.

Therefore, the independence system is bit-concave, and this shows the existence of a randomized strategy with robustness $1/\ln(4)$ by [8]. □

We note that this proposition answers a question posed by [8]:

> *We are not aware of natural systems that are bit-concave but not concave.*

Indeed, an independence system defined by an instance (E, p, w, C) of the knapsack problem with unit densities is bit-concave by Proposition 3.1. On the other hand, such an independence system is not necessarily concave, i.e., $2p(\mathrm{OPT}_{k+1}) \geq p(\mathrm{OPT}_k) + p(\mathrm{OPT}_{k+2})$ does not necessarily hold when p is not a bit-function. To see this, consider the instance of the knapsack problem such that there are four items, the values of $p_e = w_e$ are $5, 2, 2$, and 2, respectively, and $C = 6$. In this instance, $p(\mathrm{OPT}_1) = p(\mathrm{OPT}_2) = 5$ and $p(\mathrm{OPT}_3) = 6$, which shows that $2p(\mathrm{OPT}_2) < p(\mathrm{OPT}_1) + p(\mathrm{OPT}_3)$.

4 Extension to General Independence Systems

In this section, we extend our results to general independence systems. We show positive and negative results in Sections 4.1 and 4.2, respectively.

4.1 $\Omega(1/\log \rho(\mathcal{F}))$-robustness As we have already seen in Theorem 3.1, Strategy 1 is $\Omega(1/\log \rho(\mathcal{F}))$-robust if the independence system is defined by the knapsack problem. This result is extended to general independence systems.

Theorem 4.1. *For an arbitrary independence system $(E, \mathcal{F})$, there is a randomized strategy with robustness $\Omega(1/\log \rho(\mathcal{F}))$.*

The proof is the same as Theorem 3.1. That is, our randomized strategy is described as follows, where OPT_k is an optimal feasible set subject to $|\mathrm{OPT}_k| \leq k$ and $m = \lceil \log \rho(\mathcal{F}) \rceil$.

Strategy 3. Choose $X_i := \mathrm{OPT}_{2^i a_{\min}}$ with probability $1/(m+1)$ for each $i \in \{0, 1, \ldots, m\}$.

Furthermore, if OPT_k is (approximately) computable in polynomial time, then Strategy 3 is obtained in polynomial time.

4.2 Upper bounds on robustness

In this subsection, we show hardness in general independence systems. More precisely, we improve the upper bounds given in Theorem 2.2 to $\mathrm{O}(1/\log \mu(\mathcal{F}))$ and $\mathrm{O}(1/\log \rho(\mathcal{F}))$ for general independence systems.

Theorem 4.2. *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.

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

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

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
0	M^{2T}	1
1	M^{2T-1}	M^2
2	M^{2T-2}	M^4
	$\vdots$	
i	M^{2T-i}	M^{2i}
	$\vdots$	
$T-1$	M^{T+1}	M^{2T-2}
T	M^T	M^{2T}

for $i = 0, 1, \ldots, T$, 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 4.2 shows that the robustness $\Omega(1/\log \rho(\mathcal{F}))$ given in Theorem 4.1 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 3.2, 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.

Acknowledgements

The authors thank for the anonymous reviewers for their valuable comments.

References

[1] A. Caprara, H. Kellerer, U. Pferschy, and D. Pisinger, *Approximation algorithms for knapsack problems with cardinality constraints*, Eur. J. Oper. Res., 123 (2000), pp. 333–345.

[2] R. Fujita, Y. Kobayashi, and K. Makino, *Robust matchings and matroid intersections*, SIAM J. Discrete Math., 27 (2013), pp. 1234–1256.

[3] R. Hassin and S. Rubinstein, *Robust matchings*, SIAM J. Discrete Math., 15 (2002), pp 530–537.

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

[5] N. Kakimura and K. Makino, *Robust independence systems*, SIAM J. Discrete Math., 27 (2013), pp. 1257–1273.

[6] N. Kakimura, K. Makino, and K. Seimi, *Computing knapsack solutions with cardinality robustness*, Japan J. Indust. Appl. Math., 29 (2012), pp. 469–483.

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

[8] J. Matuschke, M. Skutella, and J.A. Soto, *Robust randomized matchings*, in P. Indyk, ed., Proc. 26th Ann. ACM-SIAM Symp. Discrete Algorithms, 2015, pp. 1904–1915.

[9] J. Mestre, *Greedy in approximation algorithms*, in Y. Azar and T. Erlebach, eds., Proc. 14th Ann. Eur. Symp. Algorithms, LNCS 4168, Springer-Verlag, 2006, pp. 528–539.