Stochastic Covering and Adaptivity*

Michel Goemans[1] and Jan Vondrák[2]

[1] Department of Mathematics, MIT, Cambridge, MA 02139, USA, goemans@math.mit.edu
[2] Microsoft Research, Redmond, WA 98052, USA, vondrak@microsoft.com

Abstract. We introduce a class of "stochastic covering" problems where the target set X to be covered is fixed, while the "items" used in the covering are characterized by probability distributions over subsets of X. This is a natural counterpart to the stochastic packing problems introduced in [5]. In analogy to [5], we study both adaptive and non-adaptive strategies to find a feasible solution, and in particular the *adaptivity gap*, introduced in [4].
It turns out that in contrast to deterministic covering problems, it makes a substantial difference whether items can be used repeatedly or not. In the model of Stochastic Covering with item multiplicity, we show that the worst case adaptivity gap is $\Theta(\log d)$, where d is the size of the target set to be covered, and this is also the best approximation factor we can achieve algorithmically. In the model without item multiplicity, we show that the adaptivity gap for Stochastic Set Cover can be $\Omega(d)$. On the other hand, we show that the adaptivity gap is bounded by $O(d^2)$, by exhibiting an $O(d^2)$-approximation non-adaptive algorithm.

1 Introduction

Stochastic optimization deals with problems involving uncertainty on the input. We consider a setting with multiple stages of building a feasible solution. Initially, only some information about the probability distribution of the input is available. At each stage, an "item" is chosen to be included in the solution and the precise properties of the item are revealed (or "instantiated") after we commit to selecting the item irrevocably. The goal is to optimize the expected value/cost of the solution. This model follows the framework of Stochastic Packing [4, 5] where the problem is to select a set of items with random sizes, satisfying given capacity contraints. We obtain profit only for those items that fit within the capacity; as soon as a capacity constraint is violated, the procedure terminates and we do not receive any further profit. It is an essential property of this model that once an item is selected, it cannot be removed from the solution. The objective is to maximize the expected profit obtained.

In this paper, we study a class of problems in a sense dual to Stochastic Packing: Stochastic Covering. Here, items come with random sizes again but the

* Research Supported in part by NSF grants CCR-0098018 and ITR-0121495, and ONR grant N00014-05-1-0148.

goal is to select sufficiently many items so that given covering constraints are satisfied. An example is the Stochastic Set Cover problem where the ground set X is fixed, while the items are characterized by probability distributions over subsets of X. We select items knowing only these distributions. Each item turns out to be a certain subset of X and we repeat this process until the entire set X is covered. For each item used in the solution, we have to pay a certain cost. The objective is then to minimize the expected cost of our solution.

A central paradigm in this setting is the notion of *adaptivity*. Knowing the instantiated properties of items selected so far, we can make further decisions based on this information. We call such an approach an *adaptive policy*. Alternatively, we can consider a model where this information is not available and we must make all decisions beforehand. This means, we choose an ordering of items to be selected, until a feasible solution is obtained, only based on the known probability distributions. Such an approach is called a *non-adaptive policy*. A fundamental question is, what is the benefit of being adaptive? We measure this benefit quantitatively as the ratio of expected costs incurred by optimal adaptive vs. non-adaptive policies (*the adaptivity gap*). A further question is whether a good adaptive or non-adaptive policy can be found efficiently.

1.1 Definitions

Now we define the class of problems we are interested in. The input comes in the form of a collection of *items*. Item i has a scalar value v_i and a vector size $\boldsymbol{S}(i)$. Unless otherwise noted, we assume that $\boldsymbol{S}(i)$ is a random vector with nonnegative components, while v_i is a deterministic nonnegative number. The random size vectors of different items are assumed independent.

We start with the deterministic form of a general covering problem, which is known under the name of a *Covering Integer Program* (see [17]). The forefather of these problems is the well-known Set Cover.

Definition 1 (CIP). *Given a collection of sets $\mathcal{F} = \{S(1), S(2), \ldots, S(n)\}$, $\bigcup_{i=1}^{n} S(i) = X$, Set Cover is the problem of selecting as few sets as possible so that their union is equal to X.*

More generally, given a nonnegative matrix $A \in \mathbb{R}_+^{n \times d}$ and vectors $\boldsymbol{b} \in \mathbb{R}_+^d, \boldsymbol{v} \in \mathbb{R}_+^n$, a Covering Integer Program (CIP) is the problem of minimizing $\boldsymbol{v} \cdot \boldsymbol{x}$ subject to $A\boldsymbol{x} \geq \boldsymbol{b}$ and $\boldsymbol{x} \in \{0,1\}^d$. Set Cover corresponds to the case where A is a $0/1$ matrix, with columns representing the sets $S(1), \ldots, S(n)$.

We define a Stochastic Covering problem as a stochastic variant of CIP where the columns of A, representing items sizes $\boldsymbol{S}(i)$, are random. The "demand vector" $\boldsymbol{b}$ is considered deterministic. By scaling, we can assume that $\boldsymbol{b} = \mathbf{1} = (1, 1, \ldots, 1)$.

Definition 2 (Stochastic Covering). *Stochastic Covering (SP) is a stochastic variant of a CIP where A is a random matrix whose columns are independent random nonnegative vectors, denoted $\boldsymbol{S}(1), \ldots \boldsymbol{S}(n)$. Stochastic Set Cover is a* *special case where the $\boldsymbol{S}(i)$ are random 0/1 vectors. A given instantiation of a set of items F is feasible if $\sum_{i \in F} \boldsymbol{S}(i) \geq \mathbf{1}$. The value of $\boldsymbol{S}(i)$ is instantiated and fixed once we include the item i in F. Once this decision is made, the item cannot be removed.*

When we refer to a stochastic optimization problem "with multiplicity", it means that each item on the input comes with an unlimited number of identical copies. This makes sense for deterministic CIP as well, where we could allow arbitrary integer vectors $\boldsymbol{x} \in \mathbb{Z}_+^n$. In the stochastic case, this means that the probability distributions for the copies of each item are identical; their instantiated sizes are still independent random variables.

We require a technical condition that the set of all items is feasible with probability 1. For Stochastic Covering with multiplicity, it is sufficient to require that the set of all items is feasible with positive probability.

For all variants of Stochastic Covering problems, we consider adaptive and non-adaptive policies.

Definition 3 (Adaptive and non-adaptive policies.). *A non-adaptive policy is a fixed ordering of items to be inserted.*

An adaptive policy is a function $\mathcal{P} : 2^{[n]} \times \mathbb{R}_+^d \to [n]$. The interpretation of $\mathcal{P}$ is that given a configuration $(\mathcal{I}, \boldsymbol{b})$ where $\mathcal{I}$ represents the items still available and $\boldsymbol{b}$ the remaining demand, $\mathcal{P}(\mathcal{I}, \boldsymbol{b})$ determines which item should be chosen next among the items in $\mathcal{I}$.

The cost incurred by a policy is the total value of the items used until a feasible covering is found. For an instance of a Stochastic Covering problem, we define

- *$ADAPT$ = minimum expected cost incurred by an adaptive policy.*
- *$NONADAPT$ = minimum expected cost incurred by a non-adaptive policy.*
- *$\alpha = NONADAPT/ADAPT$ is the adaptivity gap.*

For a class of Stochastic Covering problems, we define α^* as the maximum possible adaptivity gap.

1.2 Our results

We present several results on Stochastic Covering problems. We develop an LP bound on the adaptive optimum, based on the notion of "mean truncated size". For Stochastic Covering with multiplicity, we show that the worst case adaptivity gap is $\Theta(\log d)$. We prove the upper bound by presenting an efficient non-adaptive $O(\log d)$-approximation algorithm, based on the LP bound.

For Stochastic Covering without multiplicity, we have results in the special case of Stochastic Set Cover. We show that the adaptivity gap in this case can be $\Omega(d)$ and it is bounded by $O(d^2)$. Again, the upper bound is constructive, by an efficient non-adaptive approximation algorithm. Also, we show an adaptive $O(d)$-approximation algorithm for Stochastic Set Cover. This, however, does not bound the worst-case adaptivity gap which could be anywhere between $\Omega(d)$ and $O(d^2)$. We leave this as an open question.

1.3 Previous work

Stochastic optimization with recourse. Recently, stochastic optimization has come to the attention of the computer science community. An optimization model which has been mainly under scrutiny is the *two-stage stochastic optimization with recourse* [10, 8, 15]. In the first stage, only some information about the probability distribution of possible inputs is available. In the second stage, the precise input is known and the solution must be completed at any cost. The goal is to minimize the expected cost of the final solution. This model has been also extended to multiple stages [9, 16]. However, an essential difference between this model and ours is whether the randomness is in the properties of items forming a solution or in the demands to be satisfied. Let's illustrate this on the example of Set Cover: Shmoys and Swamy consider in [15] a Stochastic Set Cover problem where the sets to be chosen are deterministic and there is a random target set A to be covered. In contrast, we consider a Stochastic Set Cover problem where the target set is fixed but the covering sets are random. This yields a setting of a very different flavor.

Stochastic Knapsack. The first problem analyzed in our model of multi-stage optimization with adaptivity was the Stochastic Knapsack [4]. The motivation for this problem is in the area of *stochastic scheduling* where a sequence of jobs should be scheduled on a machine within a limited amount of time. The goal is to maximize the expected profit received for jobs completed before a given deadline. The jobs are processed one by one; after a job has been completed, its precise running time is revealed - but then it is too late to remove the job from the schedule. Hence the property of *irrevocable decisions*, which is essential in the definition of our stochastic model.

In [4, 6], we showed that adaptivity can provide a certain benefit which is, however, bounded by a constant factor. A non-adaptive solution which achieves expected value at least $1/4$ of the adaptive optimum is achieved by a greedy algorithm which runs in polynomial time. Thus the adaptivity gap is upper-bounded by 4. Concerning adaptive approximation, we showed that for any $\epsilon > 0$, there is a polynomial-time adaptive policy achieving at least $1/3 - \epsilon$ of the adaptive optimum.

Stochastic Packing. Stochastic Packing problems generalize the Stochastic Knapsack in the sense that we allow multidimensional packing contraints. This class contains many combinatorial problems: set packing, maximum matching, b-matching and general Packing Integer Programs (PIP, see e.g. [14]). In the stochastic variants of these problems we consider items with random vector sizes which are instantiated upon inclusion in the solution. Our motivation for this generalization is *scheduling with multiple resources*.

The analysis of Stochastic Packing in [5] reveals a curious pattern of results. Let us present it on the example of Stochastic Set Packing. Here, each item is defined by a value and a probability distribution over subsets $A \subseteq X$ where X is a ground set of cardinality $|X| = d$. A feasible solution is a collection of disjoint sets. It is known that for deterministic Set Packing, the greedy algorithm provides an $O(\sqrt{d})$-approximation, and there is a closely matching in-approximability result which states that for any fixed $\epsilon > 0$, a polynomial-time $d^{1/2-\epsilon}$-approximation algorithm would imply $NP = ZPP$ [2].

For Stochastic Set Packing, it turns out that the adaptivity gap can be as large as $\Theta(\sqrt{d})$. On the other hand, this is the worst case, since there is a polynomial-time non-adaptive policy which gives an $O(\sqrt{d})$-approximation of the *adaptive optimum*. Note that even with an adaptive policy, we cannot hope for a better approximation factor, due to the hardness result for deterministic Set Packing.

These results hint at a possible underlying connection between the quantities we are investigating: deterministic approximability, adaptivity gap and stochastic approximability. Note that there is no reference to computational efficiency in the notion of adaptivity gap, so a direct connection with the approximability factor would be quite surprising.

In this paper, we are investigating the question whether such phenomena appear in the case of covering problems as well.

Covering Integer Programs. Stochastic Covering problems can be seen as generalizations of Covering Integer Programs (CIP, see [17]). The forefather of Covering Integer Programs is the well-known Set Cover problem. For Set Cover, it was proved by Johnson [11] that the greedy algorithm gives an approximation factor of $\ln d$. This result was extended by Chvátal to the weighted case [3]. The same approximation guarantee can be obtained by a linear programming approach, as shown by Lovász [13]. Finally, it was proved by Uriel Feige [7] that these results are optimal, in the sense that a polynomial-time $(1-\epsilon)\ln d$ approximation algorithm for Set Cover would imply $NP \subset TIME(n^{O(\log\log n)})$.

Note. Usually the cardinality of the ground set is denoted by n, but to be consistent with Stochastic Packing problems, we view this parameter as "dimension" and denote it by d.

For general Covering Integer Problems, the optimal approximation has been found only recently [1, 12]. The approximation factor turns out to be again $O(\log d)$ but the approximation algorithm is more sophisticated. Also, the natural LP can have an arbitrarily large integrality gap.

2 Stochastic Covering with multiplicity

Let's start with the class of Stochastic Covering problems where each item can be used arbitrarily many times.

Lemma 1. *There are instances of Stochastic Set Cover with multiplicity where the adaptivity gap is at least* $0.45 \ln d$.

Proof. Consider item types for $i = 1, 2, \ldots, d$ where $\boldsymbol{S}(i) = \mathbf{0}$ or $\boldsymbol{e}_i$ with probability 1/2. All items have unit values. An adaptive policy inserts an expected number of 2 items of each type until the respective component is covered; $ADAPT \leq 2d$.

Assume that a nonadaptive policy at some point has inserted k_i items of type i, for each i. Denote the total size at that point by $\boldsymbol{S}$. We estimate the probability that this is a feasible solution:

$$\Pr[\boldsymbol{S} \geq \mathbf{1}] = \prod_{i=1}^{d} \Pr[S_i \geq 1] = \prod_{i=1}^{d} (1 - 2^{-k_i}) \leq \exp\left(-\sum_{i=1}^{d} 2^{-k_i}\right).$$

Assume that $k = \sum_{i=1}^{d} k_i = d \log_2 d$. By convexity, the probability of covering is maximized for a given d when $k_i = k/d = \log_2 d$, and then still $\Pr[\boldsymbol{S} \geq \mathbf{1}] \leq 1/e$. Thus whatever the non-adaptive policy does, there is probability $1 - 1/e$ that it needs more than $d \log_2 d$ items, which means $NONADAPT \geq (1-1/e) d \log_2 d > 0.9 d \ln d$.

Now we would like to prove that $O(\log d)$ is indeed the worst possible adaptivity gap, not only for Stochastic Set Cover with multiplicity but for all Stochastic Covering problems with multiplicity. First, we need a bound on the adaptive optimum. For this purpose, we define the *mean truncated size* of an item, in analogy to [4].

Definition 4. *We define the mean truncated size of an item with random size* $\boldsymbol{S}$ *by components as*

$$\mu_j = \mathbf{E}[\min\{S_j, 1\}].$$

For a set of items $\mathcal{A}$*, we let* $\mu_j(\mathcal{A}) = \sum_{i \in \mathcal{A}} \mu_j(i)$.

We prove that in expectation, the mean truncated size of the items inserted by any policy must be at least the demand required for each coordinate.

Lemma 2. *For a Stochastic Covering problem and any adaptive policy, let* $\mathcal{A}$ *denote the (random) set of items which the policy uses to achieve a feasible covering. Then for each component* j*,*

$$\mathbf{E}[\mu_j(\mathcal{A})] \geq 1.$$

Proof. Consider component j. Denote by $M_j(c)$ the minimum expected $\mu_j(\mathcal{A})$ for a set $\mathcal{A}$ that an adaptive policy needs to insert in order to satisfy remaining demand c in the j-th component. We prove, by induction on the number of available items, that $M_j(c) \geq c$. Suppose that an optimal adaptive policy, given remaining demand c, inserts item i. Denote by $cover(i, c)$ the indicator variable of the event that item i covers the remaining demand (i.e., $S_j(i) \geq c$, and in that case the policy terminates). We denote by $\tilde{s}_j(i)$ the truncated size $\tilde{s}_j(i) = \min\{S_j(i), 1\}$:

$$M_j(c) = \mu_j(i) + \mathbf{E}[(1 - cover(i,c)) M_j(c - \tilde{s}_j(i))] = \mathbf{E}[\tilde{s}_j(i) + (1 - cover(i,c)) M_j(c - \tilde{s}_j(i))]$$

and using the induction hypothesis,

$$M_j(c) \geq \mathbf{E}[\tilde{s}_j(i) + (1 - cover(i,c))(c - \tilde{s}_j(i))] = c - \mathbf{E}[cover(i,c)(c - \tilde{s}_j(i))] \geq c$$

since $cover(i,c) = 1$ only if $\tilde{s}_j(i) \geq c$.

Note that this lemma does not depend on whether multiplicity of items is allowed or not. In any case, having this bound, we can write an LP bounding the expected cost of the optimal adaptive policy. We introduce a variable x_i for every item i which can be interpreted as the probability that a policy ever inserts item i. For problems with item multiplicities, x_i represents the expected number of copies of item i inserted by a policy. Then the expected cost of the policy can be written as $\sum_i v_i x_i$. Due to Lemma 2, we know that $\mathbf{E}[\mu_j(\mathcal{A})] = \sum_i x_i \mu_j(i) \geq 1$ for any policy. So we get the following lower bound on the expected cost of any adaptive policy.

Theorem 1. *For an instance of Stochastic Covering with multiplicity, $ADAPT \geq LP$ where*

$$LP = \min \left\{ \sum_i x_i v_i : \begin{array}{l} \sum_i x_i \boldsymbol{\mu}(i) \geq \mathbf{1} \\ x_i \geq 0 \end{array} \right\}.$$

For a problem without multiplicity, $x_i \geq 0$ would be replaced by $x_i \in [0,1]$.

Now we are ready to prove an upper bound on the adaptivity gap, for $d \geq 2$. (The case $d = 1$ can be viewed as a special case of $d = 2$.)

Theorem 2. *For Stochastic Covering with multiplicity in dimension $d \geq 2$,*

$$NONADAPT \leq 12 \ln d \; ADAPT$$

and the corresponding non-adaptive policy can be found in polynomial time.

Proof. Consider the LP formulation of Stochastic Covering with multiplicity:

$$LP = \min \left\{ \sum_i x_i v_i : \sum_i x_i \boldsymbol{\mu}(i) \geq \mathbf{1}, \; x_i \geq 0 \right\}.$$

We know from Theorem 1 that $ADAPT \geq LP$. Let x_i be an optimal solution. We inflate the solution by a factor of $c \ln d$ (hence the need to be able to repeat items) and we build a random multiset $\mathcal{F}$ where item i has an expected number of copies $y_i = x_i \, c \ln d$. This can be done for example by including $\lfloor y_i \rfloor$ copies of item i deterministically and another copy with probability $y_i - \lfloor y_i \rfloor$. Then the total size of set $\mathcal{F}$ in component j can be seen as a sum of independent random $[0,1]$ variables and the expected total size is

$$\mathbf{E}[S_j(\mathcal{F})] = \sum_i y_i \mathbf{E}[S_j(i)] \geq \sum_i y_i \mu_j(i) \geq c \ln d.$$

By Chernoff bound, with $\mu = \mathbf{E}[S_j(\mathcal{F})] \geq c \ln d$ and $\delta = 1 - 1/\mu$:

$$\Pr[S_j(\mathcal{F}) < 1] = \Pr[S_j(\mathcal{F}) < (1-\delta)\mu] < e^{-\mu\delta^2/2} \leq e^{-\mu/2+1} \leq \frac{e}{d^{c/2}}.$$

We choose $c = 9$ and then by the union bound

$$\Pr[\exists j; S_j(\mathcal{F}) < 1] < \frac{e}{d^{3.5}}.$$

For $d \geq 2$, $\mathcal{F}$ is a feasible solution with a constant nonzero probability at least $1 - e/2^{3.5}$. Its expected cost is

$$\mathbf{E}[v(\mathcal{F})] = \sum_i y_i v_i = 9 \ln d \; LP \leq 9 \ln d \; ADAPT.$$

If $\mathcal{F}$ is not a feasible solution, we repeat; the expected number of iterations is $1/(1 - e/2^{3.5}) < 4/3$. Therefore

$$NONADAPT \leq 12 \ln d \; ADAPT.$$

This randomized rounding algorithm can be derandomized using pessimistic estimators in the usual way.

3 General Stochastic Covering

Now we turn to the most general class of Stochastic Covering problems, where each item can be used only once (unless the input itself contains multiple copies of it) and the random item sizes are without any restrictions. Unfortunately, in this setting there is little that we are able to do. We can write a linear program bounding the adaptive optimum, analogous to Theorem 1:

$$LP = \min \left\{ \sum_i x_i v_i : \begin{array}{l} \sum_i x_i \boldsymbol{\mu}(i) \geq \mathbf{1} \\ x_i \in [0, 1] \end{array} \right\}.$$

However, this LP can be far away from the actual adaptive optimum, even for $d = 1$. Consider one item of size $S(1) = 1 - \epsilon$ and an unlimited supply of items of size $S(2) = 1$ with probability ϵ and $S(2) = 0$ with probability $1 - \epsilon$. I.e., $\mu(1) = 1 - \epsilon$, $\mu(2) = \epsilon$. All items have unit values. We can set $x_1 = x_2 = 1$ and this gives a feasible solution with $LP = 2$. However, in the actual solution the item of size $1 - \epsilon$ does not help; we need to wait for an item of the second type to achieve size 1. This will take an expected number of $1/\epsilon$ items, therefore $ADAPT = 1/\epsilon$.

This example illustrates a more general issue with any approach using mean truncated sizes. As long as we do not use other information about the probability distribution, we would not distinguish between the above instance and one where the actual sizes of items are $\mu(1) = 1 - \epsilon$ and $\mu(2) = \epsilon$. Such an instance would indeed have a solution of cost 2. Thus using only mean truncated sizes, we cannot prove any approximation result in this case. It would be necessary to use a more complete description of the distributions of $S(i)$, but we leave this question outside the scope of this paper.

4 Stochastic Set Cover

Perhaps the circumstances are more benign in the case of Set Cover, i.e. 0/1 size vectors. However, the adaptivity gap is certainly not bounded by $O(\log d)$, when items cannot be used repeatedly.

Lemma 3. *For Stochastic Set Cover without multiplicity, the adaptivity gap can be $d/2$.*

Proof. Consider $\boldsymbol{S}(0) = \mathbf{1} - \boldsymbol{e}_k$, where $k \in \{1, 2, \ldots, d\}$ is uniformly random, $v_0 = 0$, and $\boldsymbol{S}(i) = \boldsymbol{e}_i$ deterministic, $v_i = 1$, for $i = 1, 2, \ldots, d$. An adaptive policy inserts item 0 first; assume its size is $\boldsymbol{S}(0) = \mathbf{1} - \boldsymbol{e}_k$. Then we insert item k which completes the covering for a cost equal to 1. An optimal non-adaptive policy still inserts item 0 first, but then, for any ordering of the remaining items that it chooses, the expected cost incurred before it hits the one which is needed to complete the covering is $d/2$.

The question is whether this is the worst possible example. First, let's consider the problem in dimension 1, where the size of each item is just a Bernoulli random variable. Thus the instance is completely characterized by the mean size values. In this case, a greedy algorithm yields the *optimal solution.*

Lemma 4. *For Stochastic Set Cover of one element ($d = 1$), assume the items are ordered, so that*

$$\frac{v_1}{\mu(1)} \leq \frac{v_2}{\mu(2)} \leq \frac{v_3}{\mu(3)} \leq \ldots \frac{v_n}{\mu(n)}$$

(call such an ordering "greedy"). Then inserting items in a greedy ordering yields a covering of minimum expected cost. The adaptivity gap in this case is equal to 1.

Proof. First, note that in this setting, adaptivity cannot bring any advantage. Until a feasible solution is obtained, we know that all items must have had size 0. An adaptive policy has no additional information and there is only one possible configuration for every subset of available items. Thus there is an optimal item to choose for each subset of available items and an optimal adaptive policy is in fact a fixed ordering of items.

For now, we assume that the items are not ordered and we consider any ordering (not necessarily the greedy ordering), say $(1, 2, 3, \ldots)$. The expected cost of a feasible solution found by inserting in this order is

$$C = \sum_{k=1}^{n} v_k \prod_{j=1}^{k-1} (1 - \mu(j)).$$

Let's analyze how switching two adjacent items affects C. Note that switching i and $i+1$ affects only the contributions of these two items - the terms corresponding to $k < i$ and $k > i+1$ remain unchanged. The difference in expected cost will be

$$\Delta C = v_i \left(\prod_{j=1}^{i-1} (1 - \mu(j)) \right) (1 - \mu(i+1)) \; + \; v_{i+1} \left(\prod_{j=1}^{i-1} (1 - \mu(j)) \right)$$

$$- v_i \left(\prod_{j=1}^{i-1} (1-\mu(j)) \right) - v_{i+1} \left(\prod_{j=1}^{i} (1-\mu(j)) \right)$$
$$= \left(\mu(i)\mu(i+1) \prod_{j=1}^{i-1} (1-\mu(j)) \right) \left(\frac{v_{i+1}}{\mu(i+1)} - \frac{v_i}{\mu(i)} \right).$$

Therefore, we can switch any pair of elements such that $\frac{v_i}{\mu(i)} \geq \frac{v_{i+1}}{\mu(i+1)}$ and obtain an ordering whose expected cost has not increased.

Assume that $\mathcal{O}$ is an arbitrary greedy ordering and $\mathcal{O}^*$ is a (possibly different) optimal ordering. If $\mathcal{O} \neq \mathcal{O}^*$, there must be a pair of adjacent items in $\mathcal{O}^*$ which are swapped in $\mathcal{O}$. By switching these two items, we obtain another optimal ordering $\mathcal{O}'$. We repeat this procedure, until we obtain $\mathcal{O}$ which must be also optimal.

The adaptive greedy algorithm. For Stochastic Set Cover in dimension d, we generalize the greedy algorithm in the following way: For each component $i \in [d]$, we find an optimal ordering restricted only to component i; we denote this by $\mathcal{O}_i$. Then our greedy adaptive algorithm chooses at any point a component j which has not been covered yet, and inserts the next available item from $\mathcal{O}_j$. Observe that this algorithm is *adaptive* as the decision is based on which components have not been covered yet.

Corollary 1. *For Stochastic Set Cover in dimension d, the greedy adaptive policy achieves expected cost*

$$GREEDY \leq d \cdot ADAPT.$$

Proof. When the policy chooses an item from $\mathcal{O}_j$, we charge its cost to a random variable X_j. Note that items from $\mathcal{O}_j$ can be also charged to other variables but an item which is charged to X_j can be inserted only after all items preceding it in $\mathcal{O}_j$ have been inserted already. Thus the value of X_j is at most the cost of covering component j, using the corresponding greedy ordering, and $\mathbf{E}[X_j] \leq ADAPT$. Consequently, $GREEDY = \sum_{i=1}^{d} \mathbf{E}[X_i] \leq d \cdot ADAPT$.

Thus we have a d-approximation, but this approximation algorithm is *adaptive* so it doesn't settle the adaptivity gap for Stochastic Set Packing. The final answer is unknown. The best upper bound we can prove is the following.

Theorem 3. *For Stochastic Set Cover,*

$$NONADAPT \leq d^2 \cdot ADAPT$$

and the corresponding non-adaptive policy can be found in polynomial time.

Proof. Consider the greedy ordering $\mathcal{O}_j$ for each component j. We interleave $\mathcal{O}_1, \mathcal{O}_2, \ldots, \mathcal{O}_d$ in the following way: We construct a single sequence $\mathcal{O}^* =$

$(i(1), i(2), i(3), \ldots)$ where $i(t)$ is chosen as the next available item from $\mathcal{O}_{j(t)}$; $j(t)$ to be defined. We set $X_i(0) = 0$ for each $1 \leq i \leq d$, $X_{j(t)}(t) = X_{j(t)}(t-1) + v_{i(t)}$ and $X_k(t) = X_k(t-1)$ for $k \neq j(t)$. In other words, we charge the cost of $i(t)$ to $X_{j(t)}$ which denotes the "cumulative cost" of component $j(t)$. At each time t, we choose the index $j(t)$ in order to minimize $X_{j(t)}(t)$ among all possible choices of $j(t)$.

Consider a fixed component k and the time τ when component k is covered. This is not necessarily by an item chosen from $\mathcal{O}_k$, i.e. $j(\tau)$ doesn't have to be k. If $j(\tau) = k$, denote by q the item from $\mathcal{O}_k$ covering component k: $q = i(\tau)$. If $j(\tau) \neq k$, denote by q the next item to be chosen from $\mathcal{O}_k$ if component k had not been covered yet. We denote by $\mathcal{Q}_k$ the prefix of sequence $\mathcal{O}_k$ up to (and including) item q. We claim that for any j, $X_j(\tau)$ is at most the cost of $\mathcal{Q}_k$: If there is some $X_j(\tau) > v(\mathcal{Q}_k)$, the last item that we charged to X_j should not have been chosen; we should have chosen an item from $\mathcal{O}_k$ which would have kept X_k still bounded by $v(\mathcal{Q}_k)$ and thus smaller than $X_j(\tau)$. Therefore $X_j(\tau) \leq v(\mathcal{Q}_k)$. For the total cost Z_k spent up to time τ when component k is covered, we get

$$Z_k = \sum_{j=1}^{d} X_j(\tau) \leq d \; v(\mathcal{Q}_k).$$

Now consider the set of items $\mathcal{Q}_k$ which is a prefix of $\mathcal{O}_k$. The probability that $\mathcal{Q}_k$ has length at least l is at most the probability that an (optimal) policy covering component k using the ordering $\mathcal{O}_k$ needs to insert at least l items from $\mathcal{O}_k$; this is the probability that the first $l-1$ items in $\mathcal{O}_k$ attain size 0 in component k. If $ADAPT_k$ denotes the minimum expected cost of an adaptive policy covering component k, we get $\mathbf{E}[v(\mathcal{Q}_k)] \leq ADAPT_k \leq ADAPT$ and $\mathbf{E}[Z_k] \leq d \cdot ADAPT$.

Finally, the total cost spent by our policy is $Z = \max_k Z_k$, since we have to wait for the last component to be covered. Therefore,

$$\mathbf{E}[Z] = \mathbf{E}[\max_{1 \leq k \leq d} Z_k] \leq \sum_{k=1}^{d} \mathbf{E}[Z_k] \leq d^2 \; ADAPT.$$

5 Concluding remarks

We have seen that allowing or not allowing items to be used repeatedly makes a significant difference in Stochastic Covering. The case where items can be used repeatedly is basically solved, with the worst-case adaptivity gap and polynomial-time approximation factor being both on the order of $\Theta(\log d)$. This would support the conjecture that there is some connection between the adaptivity gap and the optimal approximation factor for the deterministic problem. However, the general class of Stochastic Covering problems without item multiplicity does not follow this pattern. The adaptivity gap for Set Cover can be $\Omega(d)$, while the optimal approximation in the deterministic case, as well as the integrality gap of the associated LP, is $O(\log d)$.

Our main open question is what is the worst possible adaptivity gap for Set Cover. We conjecture that it is $\Theta(d)$ but we are unable to prove this. Also, it remains to be seen what can be done for general Stochastic Covering when the complete probability distributions of item sizes are taken into account.

References

1. Robert D. Carr, Lisa K. Fleischer, Vitus J. Leung, and Cynthia A. Phillips. Strengthening integrality gaps for capacitated network design and covering problems. In *SODA*, pages 106–115, 2000.
2. C. Chekuri and S. Khanna. On multidimensional packing problems. *SIAM J. Computing*, 33(4):837–851, 2004.
3. V. Chvátal. A greedy heuristic for the set covering problem. *Math. Oper. Res*, 4:233–235, 1979.
4. B. Dean, M. X. Goemans, and J. Vondrák. Approximating the stochastic knapsack: the benefit of adaptivity. In *FOCS*, pages 208–217, 2004.
5. B. Dean, M. X. Goemans, and J. Vondrák. Adaptivity and approximation for stochastic packing problems. In *SODA*, pages 395–404, 2005.
6. B. Dean, M. X. Goemans, and J. Vondrák. Approximating the stochastic knapsack: the benefit of adaptivity. *Journal version, submitted*, 2005.
7. U. Feige. A threshold of ln n for approximating set cover. *JACM*, 45(4):634–652, 1998.
8. A. Gupta, M. Pál, R. Ravi, and A. Sinha. Boosted sampling: approximation algorithms for stochastic optimization. In *SODA*, pages 417–426, 2004.
9. A. Gupta, M. Pál, R. Ravi, and A. Sinha. What about wednesday? approximation algorithms for multistage stochastic optimization. In *APPROX*, pages 86–98, 2005.
10. N. Immorlica, D. Karger, M. Minkoff, and V. Mirrokni. On the costs and benefits of procrastination: Approximation algorithms for stochastic combinatorial optimization problems. In *SODA*, pages 184–693, 2004.
11. D. Johnson. Approximation algorithms for combinatorial problems. *J. Comp. Syst. Sci.*, 9:256–216, 1974.
12. S. Kolliopoulos and N. Young. Tight approximation results for general covering integer programs. In *FOCS*, pages 522–531, 2001.
13. L. Lovász. On the ratio of the optimal integral and fractional covers. *Disc. Math.*, 13:256–278, 1975.
14. P. Raghavan. Probabilistic construction of deterministic algorithms: approximating packing integer programs. *J. Comp. and System Sci.*, 37:130–143, 1988.
15. D. Shmoys and C. Swamy. Stochastic optimization is (almost) as easy as deterministic optimization. In *FOCS*, pages 228–237, 2004.
16. D. Shmoys and C. Swamy. Sampling-based approximation algorithms for multistage stochastic optimization. In *FOCS*, pages 357–366, 2005.
17. A. Srinivasan. Improved approximations of packing and covering problems. In *STOC*, pages 268–276, 1995.