Approximation Algorithms for Stochastic and Risk-Averse Optimization*

Aravind Srinivasan†

Abstract

We present improved approximation algorithms in stochastic optimization. We prove that the multi-stage stochastic versions of covering integer programs (such as set cover and vertex cover) admit essentially the same approximation algorithms as their standard (non-stochastic) counterparts; this improves upon work of Swamy & Shmoys that shows an approximability which depends multiplicatively on the number of stages. We also present approximation algorithms for facility location and some of its variants in the 2-stage recourse model, improving on previous approximation guarantees.

1 Introduction.

Stochastic optimization attempts to model uncertainty in the input data via probabilistic modeling of future information. It originated in the work of Beale [1] and Dantzig [4] five decades ago, and has found application in several areas of optimization. There has been a flurry of algorithmic activity over the last five years in this field, especially from the viewpoint of approximation algorithms. See the survey [33] for a thorough discussion of this area.

In this work, we present improved approximation algorithms for various basic problems in stochastic optimization. We start by recalling the widely-used 2-*stage recourse model* [33]. Information about the input instance is revealed in two stages here. In the first, we are given access to a distribution D over possible realizations of future data, each such realization called a *scenario*; given D, we can commit to an anticipatory part x of the total solution, which costs us $c(x)$. In the second stage, a scenario A is sampled from D and given to us, specifying the complete instance. We may then augment x by taking recourse actions y_A that cost us the additional amount of $f_A(x, y_A)$ in order to construct a feasible solution for the complete instance. The algorithmic goal is to construct x efficiently, as well as y_A efficiently (precomputed for all A if possible, or computed when A is revealed to us), in order to minimize the total expected cost, $c(x) + \mathbf{E}_A[f_A(x, y_A)]$. (In the case of randomized algorithms, we further take the expectation over the random choices of the algorithm.) This is the basic cost-model. We will also study "risk-averse" relatives of this expectation-minimization version. There is a natural extension of the above to $k \geq 2$ stages; see [32] for a nice motivating example for the case where $k > 2$, and for the precise model. We present just those details of this model that are relevant for our discussion, in § 2.

As an example, the 2-stage version of set cover is as follows. As usual, we have a finite ground set X and a given family of subsets $S_1, S_2, \ldots, S_m$ of X; the stochasticity comes from the fact that the actual set of elements to be covered could be a *subset* of X, about which we only have probabilistic information. As above, we can sample D to get an idea of this subset in stage I; we can also buy each S_j for some given cost c_j in stage I. Of course, the catch is that future costs are typically more: i.e., for all j and A, the cost $c_{A,j}$ of buying S_j under scenario A in stage II, could be much more than c_j. This reflects the increased cost of rapid-response, as opposed to the advance provisioning of any set S_j. As in set cover, a feasible solution is a collection of S_j that covers all of the finally-revealed elements A. Thus, we will choose some collection x of S_j in stage I by using the distributional information about A, and then augment x by further sets S_j in stage II when we know A. One basic goal is to minimize the total expected cost of the two stages.

How is D specified? As mentioned in [33], there has been recent work in algorithms where the data (e.g., demands) come from a product of *independent*, explicitly-given distributions (see, e.g., the discussions in [18, 13, 5]). One major advantage here is that it can succinctly capture even exponentially many scenarios. However, just as in [22, 10, 24, 11, 32, 33], we are interested in dealing with correlations that arise in the data (e.g., correlations due to geographic proximity of clients), and hence will not deal with such *independent activation* models here. So, our general definition is where we are given access to a black-box which can generate samples according to D. Alternatively, we could be explicitly given the list of scenarios and their

*Supported in part by NSF Award CCR-0208005 and NSF ITR Award CNS-0426683.

†Department of Computer Science and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742, USA.

respective probabilities. In this case, algorithms that run in time polynomial in the other input parameters naturally require that the total number of scenarios be polynomially bounded. A natural question to ask is: can we "reduce" the former model to the latter, by taking some polynomial number of scenarios from the black box, and constructing an explicit list of scenarios using their empirical probabilities? Indeed, this *sample-average approximation* method is widely used in practice: see, e.g., [16, 34]. The work of [24, 2, 26] has shown that we can indeed reduce the black-box model to the polynomial-scenario model for the case of $k = 2$ stages, by a careful usage of sampling for the problems we study here: the error introduced by the sampling will translate to a multiplicative $(1 + \epsilon)$ factor in the approximation guarantee, where ϵ can be made as small as any desired inverse-polynomial of the input size. Therefore, we will define the k-stage model using the details relevant to us in § 2; in § 3, where we only deal with $k = 2$ stages, we will assume the polynomial-scenario model.

Our results are as follows. We consider the k-stage model in § 2; all the problems and results here, as well as earlier work on these problems, is for arbitrary constants k. The *boosted sampling* approach of [11] leads to approximation guarantees that are exponential for k for problems such as vertex cover (and better approximations for the k-stage Steiner tree problem). This was improved in [32], leading to approximation guarantees for vertex cover, set cover, and facility location that are k times their standard (non-stochastic) threshold: e.g., approximation guarantees of $2k + \epsilon$ and $k \ln n$ for vertex cover and set cover respectively are developed in [32]. Removing this dependence on k is mentioned as an open problem in [33]. We resolve this by developing simple randomized approximation algorithms that yield, for the family of covering integer programs, essentially the same approximation algorithms as for their non-stochastic counterparts. In particular, we get guarantees of $2 + \epsilon$ and $(1 + o(1)) \ln n$ respectively, for vertex cover and set cover. Except for a somewhat non-standard version of vertex cover studied in [22], these are improvements even for the case of $k = 2$. Chaitanya Swamy (personal communication, June 2006) has informed us that Kamesh Munagala has independently obtained the result for k-stage vertex cover.

Our next object of study is the classical facility location problem. Recall that in the standard (non-stochastic) version of the facility location problem, we are given a set of clients $\mathcal{D}$ and a set of facilities $\mathcal{F}$. The distance from client j to facility i is c_{ij}, and these values form a metric. Given a cost f_i for opening each facility i, we want to open some of the facilities, so that the sum of the opening costs and the distances traveled by each client to its closest open facility, is minimized. (As usual, all results and approximations translate without any loss to the case where each client i has a demand $d_i \geq 0$ indicating the number of customers at i, so we just discuss the case $d_i \in \{0, 1\}$ here.) Starting with [25], there has been a steady stream of constant-factor approximation algorithms for the problem, drawing from and contributing to techniques in LP rounding, primal-dual methods, local search, greedy algorithms etc. The current-best lower and upper bounds on the problem's approximability are $1.46 \cdots$ [9] and 1.52 [17]. The stochastic version of the facility location problem has also received a good deal of attention in the case of $k = 2$ stages [22, 31, 24]. In the 2-stage stochastic version, the results of [24, 2, 26] let us assume the polynomial-scenario model as mentioned above. Here, each facility i can be opened at cost f_i^I in stage I, and at a cost f_i^A when a scenario A materializes in stage II; each scenario A is simply a subset of $\mathcal{D}$, indicating the actual set of clients that need to be served under this scenario. The goal is to open some facilities in stage I and then some in stage II; we develop improved approximation algorithms in two settings here, as discussed next.

Our first setting is the basic one of minimizing the total expected cost, just as for covering problems. That is, we aim to minimize the expected "client-connection costs + facility-opening cost", where the expectation is both over the random emergence of scenarios, and internal random choices of our algorithm. We develop a 2.369-approximation in § 3.1, improving upon the current-best value of 3 [24, 33]. Our approach here crucially exploits an asymmetry in the facility-location algorithms of [14, 15].

Our second setting involves an additional ingredient of "risk-aversion", various facets of which have been studied in works including [12, 7, 26]. The motivation here is that a user may be risk-averse, and may not want to end up paying much if (what was perceived to be) a low-probability scenario emerges: minimizing the overall expectation alone may not suffice. Therefore, as in [12], we aim for "local" algorithms: those where for each scenario A, its expected final (rounded) cost is at most some guaranteed (constant) factor times its fractional counterpart Val_A. Such a result is very useful since it allows the inclusion of budgets to the various scenarios, and to either prove that these are infeasible, or to come within a constant factor of each particular budget in expectation [12]. In § 3.2, we prove that a randomized rounding scheme of [31] actually guarantees that for each scenario A, its expected final (rounded) cost is at most $3.25 \cdot Val_A$, improving on the

$3.378 \cdot Val_A$ bound of [31]. (In the previous version of this paper, these values 3.25 and 3.378 were 3.095 and 3.225 respectively. However, as pointed out to us by Chaitanya Swamy in September 2006, there is a small error in the original thesis [30] which is fixed in [31]; the value 3.225 and the corresponding proof of [30], are revised to 3.378 with a corrected proof in [31]. This necessitates the change in our paper as well.) Our improvement comes by an improved analysis of a key inequality of [31] (bound (3.16) of this paper). The r.h.s. of this inequality is upper-bounded in [31] by bounding the two summands separately. See also [29] for a similar approach. We are able to bound this term altogether, leading to our improvement: the supporting intuition is that the bounds for the two summands from [31] are tight in two *distinct* regimes.

Thus, we present improved approximation algorithms in stochastic optimization, both for two stages and multiple stages, based on LP-rounding. Our main results are presented in Theorem 2.1, Theorem 3.1, and Theorem 3.2.

2 Multi-stage covering problems.

We show that general stochastic k-stage covering integer programs (those with all coefficients being non-negative, and with the variables x_j allowed to be arbitrary non-negative integers) admit essentially the same approximation ratios as their usual (non-stochastic) counterparts. The model is that we can "buy" each x_j in k stages: the final value of x_j is the sum of all values bought for it. We also show that k-stage vertex cover can be approximated to within $2+\epsilon$; similarly, k-stage set cover with each element of the universe contained in at most b of the given sets, can be approximated to within $b+\epsilon$.

Model: Given the results of [32], we can model a k-stage covering integer program (CIP) for our purposes as follows; viewing the problem thus gives a clean "online" framework within which to develop our algorithms. (The actual model allows the algorithm-designer further leeway, but it suffices to develop an algorithm under the constraints given next. There is a *hidden* covering problem "minimize $c^T \cdot x$ subject to $Ax \geq b$ and all variables in x being non-negative integers". For notational convenience for the set-cover problem, we let n be the number of rows of A; also, the variables in x are indexed as $x_{j,\ell}$, where $1 \leq j \leq m$ and $1 \leq \ell \leq k$. This program, as well as a feasible *fractional* solution x^* for it, are revealed to us in k stages as follows. In each stage ℓ ($1 \leq \ell \leq k$), we are given the ℓth-stage fractional values $\{x^*_{j,\ell} : 1 \leq j \leq m\}$ of the variables, along with their columns in the coefficient matrix A, and their coefficient in the objective function c. Given some values like this, we need to round them right away at stage ℓ using randomization if necessary, *irrevocably*. The goal is to develop such a rounded vector $\{y_{j,\ell} : 1 \leq j \leq m, 1 \leq \ell \leq k\}$ that satisfies the constraints $Ay \geq b$, and whose (expected) approximation ratio $c^T \cdot y / c^T \cdot x^*$ is small. Our results here are summarized as follows:

THEOREM 2.1. *We obtain randomized λ-approximation algorithms for k-stage stochastic CIPs for arbitrary fixed k, with values of λ as follows. (The running time is polynomial for any fixed k, and λ is independent of k.) (i) For general CIPs, with the linear system scaled so that all entries of the matrix A lie in $[0,1]$ and $\min_i b_i = B$, we have $\lambda = 1 + O(\min\{(\ln n)/B, \sqrt{(\ln n)/B}\})$. (ii) For set cover with element-degree (maximum number of given sets containing any element of the ground set) at most b, we have $\lambda = b + \epsilon$, where ϵ can be N^{-C} with N being the input-size and $C > 0$ being any constant. (For instance, $b = 2$ for vertex cover, where an edge can be covered only by its two end-points.)*

The "$+\epsilon$" term appears in part (ii) since the fractional solution obtained by [32] is an $(1+\epsilon)$–approximation to the actual LP. We do not mention this term in part (i), by absorbing it into the big-Oh notation. The two parts of this theorem are proved next.

2.1 A simple scheme for general CIPs.

We use our k-stage model to prove Theorem 2.1(i). We show that a simple randomized rounding approach along the lines of [21] works here: for a suitable $\lambda \geq 1$ and independently for all (j,ℓ), set $x'_{j,\ell} = \lambda \cdot x^*_{j,\ell}$, and define the rounded value $y_{j,\ell}$ to be $\lceil x'_{j,\ell} \rceil$ with probability $x'_{j,\ell} - \lfloor x'_{j,\ell} \rfloor$, and to be $\lfloor x'_{j,\ell} \rfloor$ with the complementary probability of $1-(x'_{j,\ell} - \lfloor x'_{j,\ell} \rfloor)$. Note that $\mathbf{E}[y_{j,\ell}] = x'_{j,\ell}$. We will now show that for a suitable, "not very large" choice of λ, with high probability, all constraints will be satisfied and $c^T \cdot y$ is about $\lambda \cdot (c^T \cdot x^*)$.

The proof is standard, and we will illustrate it for set cover. Note that in this case, a set of at most n elements need to be covered in the end. Set $\lambda = \ln n + \psi(n)$ for any arbitrarily slowly growing function $\psi(n)$ of n such that $\lim_{n\to\infty} \psi(n) = \infty$; run the randomized rounding scheme described in the previous paragraph. Consider any finally revealed element i, and let E_i be the event that our rounding leaves this element uncovered. Let A_i be the family of sets in the given set-cover instance that contain i; note that the fractional solution satisfies $\sum_{j \in A_i, \ell} x^*_{j,\ell} \geq 1$. Now, if $x'_{j,\ell} \geq 1$ for some pair ($j \in A_i$, ℓ), then $y_{j,\ell} \geq 1$, and so, i is

guaranteed to be covered. Otherwise,

$$\begin{aligned}\Pr[E_i] &= \prod_{j\in A_i,\ell} \Pr[y_{j,\ell}=0] \\ &= \prod_{j\in A_i,\ell} (1-x'_{j,\ell}) \\ &\le \exp(-\sum_{j\in A_i,\ell} \lambda\cdot x^*_{j,\ell}) \\ &\le \exp(-\lambda) \\ &= \exp(-\psi(n))/n = o(1/n).\end{aligned}$$

Thus, applying a union bound over the (at most n) finally-revealed elements i, we see that $\Pr[\bigwedge_i \overline{E_i}] = 1-o(1)$. So,

$$\begin{aligned}\mathbf{E}[c^T\cdot y \mid \bigwedge_i \overline{E_i}] &\le \frac{\mathbf{E}[c^T\cdot y]}{\Pr[\bigwedge_i \overline{E_i}]} \\ &= \frac{\lambda\cdot(c^T\cdot x^*)}{\Pr[\bigwedge_i \overline{E_i}]} \\ &= \frac{\lambda\cdot(c^T\cdot x^*)}{1-o(1)} \\ &= (1+o(1))\cdot\lambda\cdot(c^T\cdot x^*);\end{aligned}$$

i.e., we get an $(1+o(1))\cdot\ln n$–approximation. Alternatively, since $c^T\cdot y$ is a sum of independent random variables, we can show that it is not much more than its mean, $\lambda\cdot(c^T\cdot x^*)$, with high probability.

The analysis for general CIPs is similar; we observe that for any row i of the constraint system, $\mathbf{E}[(Ay)_i] = \lambda b_i$, use a Chernoff lower-tail bound to show that the "bad" event E_i that $(Ay)_i < b_i$ happens with probability noticeably smaller than $1/n$, and apply a union bound over all n. Choosing λ as in Theorem 2.1(i) suffices for such an analysis; see, e.g., [19]. See [6] for randomized-rounding approaches for a different model of stochastic packing problems. Also, note that there is a small probability (which may be made any desired inverse-polynomial of the input-size N) that not all constraints are satisfied by our algorithm. While such a Monte Carlo algorithm with a very small probability of failure is often sufficient, it is not clear how to completely eliminate such a possibility of failure in our stochastic setting, without losing much in the approximation guarantee. While a natural greedy approach can be used to satisfy the unsatisfied constraints in the *non-stochastic* case [28], it appears that we may need bounds on the rate by which the entries of the matrix A grow as we move to further stages, in order to bound the performance of such (say, greedy) approaches.

2.2 Vertex cover, and set cover with small degree. We now use a type of dependent rounding to prove Theorem 2.1(ii). We present the case of vertex cover ($b=2$), and then note the small modification needed for the case of general b. Note that our model becomes the following for (weighted) vertex cover. There is a hidden undirected graph $G=(V,E)$. The following happens for each vertex $v\in V$. We are revealed k fractional values $x^*_{v,1}, x^*_{v,2}, \ldots, x^*_{v,k}$ for v one-by-one, along with the corresponding weights for v (in the objective function), $c_{v,1}, c_{v,2}, \ldots, c_{v,k}$. We aim for a rounding $\{y_{v,\ell}\}$ that covers all edges in E, whose objective-function value $\sum_{\ell,v} c_{v,\ell}y_{v,\ell}$ is at most twice its fractional counterpart, $\sum_{\ell,v} c_{v,\ell}x^*_{v,\ell}$. Note that the fractional solution satisfies

$$\forall (u,v)\in E,\ (\sum_{\ell=1}^{k} x^*_{u,\ell}) + (\sum_{\ell=1}^{k} x^*_{v,\ell}) \ge 1. \tag{2.1}$$

Now, given a sequence $z=(z_1,z_2,\ldots,z_k)$ of values that lie in $[0,1]$ and arrive *online*, suppose we can define an efficient randomized procedure $\mathcal{R}$, which has the following properties:

- **(P1)** as soon as $\mathcal{R}$ gets a value z_i, it rounds it to some $Z_i\in\{0,1\}$ ($\mathcal{R}$ may use the knowledge of the values $\{z_j, Z_j : j<i\}$ in this process);
- **(P2)** $\mathbf{E}[Z_i]\le z_i$; and
- **(P3)** if $\sum_i z_i \ge 1$, then at least one Z_i is one with probability one.

Then, we can simply apply procedure $\mathcal{R}$ independently for each vertex v, to the vector $z(v)$ of *scaled* values $(\min\{2\cdot x^*_{v,1},1\}, \min\{2\cdot x^*_{v,2},1\}, \ldots, \min\{2\cdot x^*_{v,k},1\})$. Property (P2) shows that the expected value of the final solution is at most $2\cdot\sum_{\ell,v} c_{v,\ell}x^*_{v,\ell}$; also, since (2.1) shows that for any edge (u,v), at least one of the two sums $2\cdot\sum_\ell x^*_{u,\ell}$ and $2\cdot\sum_\ell x^*_{v,\ell}$ is at least 1, property (P3) guarantees that each edge (u,v) is covered with probability one. So, the only task is to define function $\mathcal{R}$.

For a sequence $z=(z_1,z_2,\ldots,z_k)$ arriving online, $\mathcal{R}$ proceeds as follows. Given z_1, it rounds z_1 to $Z_1=1$ with probability z_1, and to $Z_1=0$ with probability $1-z_1$. Next, given z_i for $i>1$:

Case I: $Z_j=1$ for some $j<i$. In this case, just set Z_i to 0.

Case II(a): $Z_j=0$ for all $j<i$, and $\sum_{\ell=1}^{i} z_\ell \ge 1$. In this case, just set Z_i to 1.

Case II(b): $Z_j=0$ for all $j<i$, and $\sum_{\ell=1}^{i} z_\ell < 1$. In this case, set $Z_i=1$ with probability $z_i/(1-\sum_{j=1}^{i-1} z_j)$, and set $Z_i=0$ with the complementary probability.

It is clear that property (P1) of $\mathcal{R}$ holds. Let us next prove property (P3). Assume that for some t,

$\sum_{i=1}^{t} z_i \geq 1$ and $\sum_{i=1}^{t-1} z_i < 1$. It suffices to prove that $\Pr[\exists i \leq t : Z_i = 1] = 1$. Note that $\Pr[\exists i \leq t : Z_i = 1]$ is the sum of $\Pr[\exists i < t : Z_i = 1]$ and $\Pr[\nexists i < t : Z_i = 1] \cdot \Pr[(Z_t = 1) \bigm| (Z_1 = Z_2 = \cdots Z_{t-1} = 0)]$, which is at least $\Pr[(Z_t = 1) \bigm| (Z_1 = Z_2 = \cdots Z_{t-1} = 0)]$, which in turn equals 1 from case II(a). This proves property (P3).

We next consider property (P2), which is immediate for $i = 1$. If there is some t such that $\sum_{i=1}^{t} z_i \geq 1$, take t to be the smallest such index; if there is no such t, define $t = k$. The required bound of (P2), $\mathbf{E}[Z_i] \leq z_i$, clearly holds for all $i > t$, since by (P3) and Case I, we have $\mathbf{E}[Z_i] = 0$ for all such i. So suppose $i \leq t$. We have from case II(a) that for all $j < t$,

$$\Pr[(Z_j = 1) \bigm| (Z_1 = Z_2 = \cdots Z_{j-1} = 0)]$$

equals

$$\frac{z_j}{1 - (z_1 + z_2 + \cdots + z_{j-1})}.$$

Note from Case I that no two Z_j can both be 1. Thus, for $1 < i \leq t$,

$$\begin{aligned}
\Pr[Z_i = 1] &= \Pr[(Z_1 = Z_2 = \cdots Z_{i-1} = 0) \wedge (Z_i = 1)] \\
&= \Pr[Z_1 = Z_2 = \cdots Z_{i-1} = 0] \cdot \\
&\quad \Pr[(Z_i = 1) \bigm| (Z_1 = Z_2 = \cdots Z_{i-1} = 0)] \\
&= (\prod_{j=1}^{i-1} (1 - \frac{z_j}{1 - (z_1 + z_2 + \cdots + z_{j-1})})) \cdot \\
&\quad \Pr[(Z_i = 1) \bigm| (Z_1 = Z_2 = \cdots Z_{i-1} = 0)] \\
&= (1 - (z_1 + z_2 + \cdots + z_{i-1})) \cdot \\
&\quad \Pr[(Z_i = 1) \bigm| (Z_1 = Z_2 = \cdots Z_{i-1} = 0)]. \qquad (2.2)
\end{aligned}$$

From Cases II(a) and II(b), $\Pr[(Z_i = 1) \bigm| (Z_1 = Z_2 = \cdots Z_{i-1} = 0)]$ is 1 if $i = t$ *and* $z_1 + z_2 + \cdots + z_t \geq 1$, and is

$$\frac{z_i}{1 - (z_1 + z_2 + \cdots + z_{i-1})}$$

otherwise; in either case, we can verify from (2.2) that $\Pr[Z_i = 1] \leq z_i$, proving (P2).

Similarly, for k-stage set cover with each element of the universe contained in at most b of the given sets, we construct $z'(v) = (\min\{b \cdot x^*_{v,1}, 1\}, \min\{b \cdot x^*_{v,2}, 1\}, \ldots, \min\{b \cdot x^*_{v,k}, 1\})$ and apply $\mathcal{R}$. By the same analysis as above, all elements are covered with probability 1, and the expected objective function value is at most $b \cdot \sum_{\ell,v} c_{v,\ell} x^*_{v,\ell}$.

Tail bounds: It is also easy to show using [20] that in addition to its expectation being at most b times the fractional value, $c^T \cdot y$ has a Chernoff-type upper bound on deviations above its mean.

3 Facility Location Problems.

We consider two variants of facility location in this section. We will consider just the case of $0-1$ demands. As usual, our algorithms directly generalize to the case of arbitrary demands with no loss in approximation guarantee. Also, as mentioned in § 1, we assume the polynomial-scenario model w.l.o.g.

3.1 Minimizing expected cost.

We study stochastic uncapacitated facility location for $k = 2$ stages, and develop a 2.369–approximation algorithm for minimizing the expected total cost. Let the set of facilities be $\mathcal{F}$, and the set of all possible clients be $\mathcal{D}$. From the results of [24, 2, 32, 26], we may assume that we are given: (i) m scenarios (indexed by A), each being a subset of $\mathcal{D}$, and (ii) an $(1+\epsilon)$-approximate solution (x, y) to the following standard LP relaxation of the problem (as in Theorem 2.1, ϵ can be made inverse-polynomially small, and will henceforth be ignored):

$$\text{minimize } \sum_{i \in \mathcal{F}} f^I_i y_i + \sum_A p_A (\sum_i f^A_i y_{A,i} + \sum_{j \in A} \sum_i c_{ij} x_{A,ij})$$

subject to:

$$\begin{aligned}
(3.3) \qquad \sum_i x_{A,ij} &\geq 1 \quad \forall A \; \forall j \in A; \\
(3.4) \qquad x_{A,ij} &\leq y_i + y_{A,i} \quad \forall i \; \forall A \; \forall j \in A; \\
x_{A,ij}, y_i, y_{A,i} &\geq 0 \quad \forall i \; \forall A \; \forall j \in A.
\end{aligned}$$

Here, f^I_i and f^A_i are the costs of opening facility i in stage I and in stage-II scenario A, respectively; c_{ij} is the cost of connecting client j to i. Each given scenario A materializes with probability p_A. Variables y_i and $y_{A,i}$ are the extents to which facility i is opened in stage I and in stage-II scenario A, respectively; $x_{A,ij}$ is the extent to which j is connected to i in scenario A.

The idea is, as in [24], to satisfy some client-scenario pairs (j, A) in Stage I; the rest will be handled in Stage II. The crucial difference is that instead of a "deterministic thresholding" to choose such Stage-I pairs as in [24], we will employ a carefully-chosen *randomized* thresholding. As we will see, this randomized scheme also fits well with a basic *asymmetry* in many known facility-location algorithms (in our case, the ones in [14, 15]).

For all i, A, and $j \in A$ such that $x_{A,ij} > 0$, write $x_{A,ij} = x^{(1)}_{A,ij} + x^{(2)}_{A,ij}$, where

$$\begin{aligned}
x^{(1)}_{A,ij} &= x_{A,ij} \cdot \frac{y_i}{y_i + y_{A,i}}; \\
x^{(2)}_{A,ij} &= x_{A,ij} \cdot \frac{y_{A,i}}{y_i + y_{A,i}}.
\end{aligned}$$

Extending this definition, if $j \in A$ and $x_{A,ij} = 0$, we define $x^{(1)}_{A,ij} = x^{(2)}_{A,ij} = 0$. Note from (3.4) that $x^{(1)}_{A,ij} \leq y_i$

and $x^{(2)}_{A,ij} \le y_{A,i}$. Let $\alpha \in (0, 1/2)$ be a constant that will be chosen later. Pick a single random real Z using the following distribution that is a mixture of continuous and discrete:

- with probability $\alpha/(1-\alpha)$, let $Z := 1/2$;
- with the complementary probability of $(1-2\alpha)/(1-\alpha)$, let Z be a random real chosen from the uniform distribution on $[\alpha, 1-\alpha]$.

The rounding for Stage I is as follows. For any pair (j, A) with $j \in A$, define $r^{(1)}_{A,j}$ (the extent to which (j, A) is satisfied in Stage I) to be

$$r^{(1)}_{A,j} = \sum_i x^{(1)}_{A,ij}; \tag{3.5}$$

the pair (j, A) is declared *selected* iff

$$Z \le r^{(1)}_{A,j}. \tag{3.6}$$

For the Stage I decisions, construct a facility-location instance $\mathcal{I}$ with each selected pair (j, A) having demand p_A and each facility i having cost f^I_i, and solve it using the approximation algorithm of [14, 15], which is described in [17] and called the *JMS Algorithm* in [17]. (The non-selected pairs are not considered in the instance $\mathcal{I}$.) In Stage II, we round separately for each scenario A as follows. Construct a facility-location instance $\mathcal{I}_A$ with a unit-demand client for each $j \in A$ such that (j, A) was *not* selected in Stage I; each facility i has cost f^A_i. Again use the JMS algorithm as described in [17] to get an approximately optimal solution for $\mathcal{I}_A$.

Analysis: It is clear that in every scenario A, we satisfy all of its demands. To analyze the expected cost of this solution (with the only randomness being in the choice of Z), we start by constructing feasible fractional solutions for the facility-location instances $\mathcal{I}$ and $\mathcal{I}_A$ (for all A). Condition on a fixed value for Z. Let us first construct a feasible fractional solution $(\hat{x}, \hat{y})$ for the stage-I instance $\mathcal{I}$: $\hat{y}_i = \min\{y_i/Z, 1\}$ for all i, and $\hat{x}_{A,ij} = x^{(1)}_{A,ij}/r^{(1)}_{A,j}$ for all selected (j, A) and all i. This is feasible since $r^{(1)}_{A,j} \ge Z$. Thus, letting $S_{j,A}$ be the indicator variable for (j, A) being selected (which is a function of Z) and recalling that each selected (j, A) has demand p_A in $\mathcal{I}$, the total "facility cost" and "connection cost" of $(\hat{x}, \hat{y})$ are

$$\sum_i \frac{y_i}{Z} \text{ and } \sum_{j,A} p_A \cdot \frac{S_{j,A}}{r^{(1)}_{A,j}} \cdot \sum_i c_{ij} x^{(1)}_{A,ij}, \tag{3.7}$$

respectively. Next consider any scenario A, and let us construct a feasible fractional solution (x', y') for $\mathcal{I}_A$. Define $r^{(2)}_{A,j} = \sum_i x^{(2)}_{A,ij}$. We may assume w.l.o.g. that equality holds in (3.3); so, $r^{(2)}_{A,j} = 1 - r^{(1)}_{A,j}$. Thus, a necessary condition for (j, A) to *not* be selected in Stage I is

$$(1 - Z) \le r^{(2)}_{A,j}. \tag{3.8}$$

This is analogous to (3.6), with Z being replaced by $1 - Z$. Thus, we can argue similarly as we did for $(\hat{x}, \hat{y})$ that $y'_i = y_{A,i}/(1-Z)$, $x'_{A,ij} = x^{(2)}_{A,ij}/r^{(2)}_{A,j}$ for all (j, A) not selected in Stage I, is a feasible fractional solution for $\mathcal{I}_A$. Since all demands here are one, the total facility cost and connection cost of (x', y') are

$$\sum_i \frac{y_{A,i}}{1-Z} \text{ and } \sum_{j,A} \frac{1 - S_{j,A}}{r^{(2)}_{A,j}} \cdot \sum_i c_{ij} x^{(2)}_{A,ij} \tag{3.9}$$

respectively.

Now, the key "asymmetry" property of the JMS algorithm is, as proven in [17], that it is a $(1.11, 1.78)$-*bifactor approximation algorithm*: given an instance for which there is a fractional solution with facility cost F and connection cost C, it produces an integral solution of cost at most $1.11F + 1.78C$. Therefore, from (3.7) and (3.9), and weighting the latter by p_A, we see that given Z, the total final cost is at most 1.11 times

$$\left[\sum_i \left(\frac{y_i}{Z} + \sum_A p_A \cdot \frac{y_{A,i}}{1-Z} \right) \right]$$

plus 1.78 times

$$\sum_{j,A} p_A \cdot \left[\left(\frac{S_{j,A}}{r^{(1)}_{A,j}} \cdot \sum_i c_{ij} x^{(1)}_{A,ij} \right) + \left(\frac{1 - S_{j,A}}{r^{(2)}_{A,j}} \cdot \sum_i c_{ij} x^{(2)}_{A,ij} \right) \right];$$

so, the expected final cost is at most the following sum of three terms:

$$\begin{aligned} &1.11 \cdot [\textstyle\sum_i (y_i \cdot \mathbf{E}[1/Z] + \sum_A p_A y_{A,i} \cdot \mathbf{E}[1/(1-Z)])] + \\ &1.78 \cdot \textstyle\sum_{j,A} p_A \cdot \frac{\mathbf{E}[S_{j,A}]}{r^{(1)}_{A,j}} \cdot \sum_i c_{ij} x^{(1)}_{A,ij} + \\ &1.78 \cdot \textstyle\sum_{j,A} p_A \cdot \frac{\mathbf{E}[1 - S_{j,A}]}{r^{(2)}_{A,j}} \cdot \sum_i c_{ij} x^{(2)}_{A,ij}. \end{aligned} \tag{3.10}$$

Since Z and $1 - Z$ have identical distributions,

$$\begin{aligned} \mathbf{E}[1/(1-Z)] &= \mathbf{E}[1/Z] \\ &= (\alpha/(1-\alpha)) \cdot 2 + \\ &\quad \frac{1-2\alpha}{1-\alpha} \cdot \frac{1}{1-2\alpha} \cdot \int_{z=\alpha}^{1-\alpha} dz/z \\ &= \frac{2\alpha + \ln((1-\alpha)/\alpha)}{1-\alpha}. \end{aligned} \tag{3.11}$$

Let us next bound $\mathbf{E}[S_{j,A}]$. Recall (3.6), and let r denote $r^{(1)}_{A,j}$. If $r < \alpha$, then $S_{j,A} = 0$; if $r \ge 1 - \alpha$, then

$S_{j,A} = 1$. Next suppose $\alpha \le r < 1/2$. Then $S_{j,A}$ can hold only if we chose to pick Z at random from $[\alpha, 1-\alpha]$, and got $Z \le r$; this happens with probability $((1-2\alpha)/(1-\alpha)) \cdot (r-\alpha)/(1-2\alpha) = (r-\alpha)/(1-\alpha) \le r/(1-\alpha)$. Finally, if $1/2 \le r < (1-\alpha)$,

$$\begin{aligned}\mathbf{E}[S_{j,A}] &= \alpha/(1-\alpha) + \\ &\quad ((1-2\alpha)/(1-\alpha)) \cdot (r-\alpha)/(1-2\alpha) \\ &= r/(1-\alpha).\end{aligned}$$

Thus, in all cases we saw here,

$$\mathbf{E}[S_{j,A}] \le r^{(1)}_{A,j}/(1-\alpha). \tag{3.12}$$

Similarly, recalling (3.8) and the fact that Z and $1-Z$ have identical distributions, we get

$$\mathbf{E}[1-S_{j,A}] \le r^{(2)}_{A,j}/(1-\alpha). \tag{3.13}$$

Plugging (3.11), (3.12), and (3.13) into (3.10) and using the fact that $x_{A,ij} = x^{(1)}_{A,ij} + x^{(2)}_{A,ij}$, we see that our expected approximation ratio is

$$\max\left\{\frac{1.78}{1-\alpha},\ \frac{1.11(2\alpha + \ln((1-\alpha)/\alpha))}{1-\alpha}\right\}.$$

A good choice of α is 0.2485, leading to an expected approximation ratio less than 2.369. Thus we get

THEOREM 3.1. *Two-stage stochastic facility location can be approximated to within a certain value ρ, whose expected value is at most* 2.369.

3.2 Facility location with per-scenario bounds. Consider again the 2-stage facility location problem, and a corresponding optimal fractional solution. We now describe a randomized rounding scheme so that for each scenario A, its expected final (rounded) cost is at most 3.25 times its fractional counterpart Val_A, improving on the $3.378 \cdot Val_A$ bound of [31].

Our algorithm is essentially the same as that of Sections 2.4 and 4.3 in [31], with an improved analysis. As in Section 2.4 of [31], we start with a randomized algorithm for standard (non-stochastic) facility location, and will then use it for stochastic facility location. So, suppose we are given a facility location instance with facilities $\mathcal{F}$ and clients $\mathcal{D}$, as well as an optimal[1] solution (x^*, y^*) to its standard LP relaxation: minimize $(\sum_{i,j} c_{ij}x_{ij} + \sum_i f_i y_i)$ subject to: (i) $\forall j$, $\sum_i x_{ij} = 1$; (ii) $\forall (i,j)$, $x_{ij} \le y_i$; (iii) $\forall (i,j)$, $x_{ij}, y_i \ge 0$. We consider the randomized algorithm of [31] to round (x^*, y^*) to a pair (X, Y); we will show

$$\mathbf{E}[\sum_i f_i Y_i] \le 1.73 \cdot \sum_i f_i y^*_i; \text{ and} \tag{3.14}$$

$$\forall j,\ \mathbf{E}[\sum_i c_{ij} X_{ij}] \le 1.73 \cdot \sum_i c_{ij} x^*_{ij}. \tag{3.15}$$

Let us present the algorithm of Section 2.4 of [31]. Let $\gamma \in (0,1)$ be a constant parameter to be chosen later. By a natural "facility cloning" approach of [3, 29], we first transform (x^*, y^*) into (x, y) such that:

(A1) $\forall j$, $\sum_i c_{ij}x_{ij} = \sum_i c_{ij}x^*_{ij}$;

(A2) $\sum_i f_i y_i = \sum_i f_i y^*_i$;

(A3) $\forall (i,j)$, $x_{ij} > 0 \Rightarrow x_{ij} = y_i$;

(A4) $\forall i$, $y_i \le \gamma$; and

(A5) Fix any $j \in \mathcal{D}$. Let $F_j = \{i : x_{ij} > 0\}$ with $|F_j| = m = m(j)$, and let π_j be some fixed permutation of F_j such that $c_{\pi_j(1)j} \le c_{\pi_j(2)j} \le \cdots \le c_{\pi_j(m)j}$. Then there exists some $v = v(j)$ such that $\sum_{i=1}^{v} x_{\pi_j(i)j} = \gamma$.

(The above transformation makes copies of some facilities, but we still refer to the new set of facilities as $\mathcal{F}$ for simplicity.)

Notation. For any j, let $v = v(j)$ be as in (A5). Define $\overline{C}_j = \sum_{i=1}^{m} c_{ij}x^*_{ij}$, $C_j(\gamma) = \sum_{i=1}^{v} c_{\pi_j(i)j}x_{\pi_j(i)j}/\gamma$, $R_j(\gamma) = c_{\pi_j(v)j}$, and $N_j = \{\pi_j(1), \pi_j(2), \ldots, \pi_j(v)\}$.

Order the clients j in non-decreasing order of $R_j(\gamma)$. Repeat the following step until no clients are left: "choose the next remaining client j in our order, call j *clustered*, and delete all remaining clients j' (including j) for which $N_{j'} \cap N_j \ne \emptyset$". At the end, if j was not declared clustered, we will call it "unclustered". Recall that Y_i is the indicator random variable for opening facility i. Independently for each clustered j, round exactly one Y_i, for $i \in N_j$, to one: for each $i \in N_j$, $\Pr[Y_i = 1] = y_i/\gamma$. (This is possible since, by (A3) and (A5), $\sum_{i \in N_j} y_i = \gamma$.) For all remaining facilities i – those that do not lie in N_j for any clustered j – independently round them such that $\Pr[Y_i = 1] = y_i/\gamma$. We then connect each j to its closest open facility; this final connection vector is denoted $X = (X_{ij})$. The following lemma is not hard to see:

LEMMA 3.1. ***([31])*** *$\mathbf{E}[\sum_i f_i Y_i] = \sum_i f_i y^*_i/\gamma$; also, for all clustered j, $\mathbf{E}[\sum_i c_{ij}X_{ij}] = C_j(\gamma) \le \overline{C}_j$.*

As in [31], the interesting case is that of *unclustered* j; our main result here is

[1] Actually an $(1+\epsilon)$-approximation, but we will ignore this by absorbing it into the approximation ratio.

LEMMA 3.2. *For all unclustered j, $\mathbf{E}[\sum_i c_{ij}X_{ij}]$ is at most*

$$\max\left\{\frac{1}{\gamma}, \frac{e^{-1}+2e^{-1/\gamma}}{1-\gamma}, 1+2e^{-1/\gamma}\right\}\cdot \overline{C}_j.$$

Proof. We provide a proof sketch. Fix any unclustered j. If $R_j(\gamma)=0$, it is easy to show that $\mathbf{E}[\sum_i c_{ij}X_{ij}] = 0$, so we will assume that $R_j(\gamma)$ is positive. Recall the definitions of F_j, m, and π_j from (A5). Fix an unclustered j, and renumber the facilities so that $\pi_j(1)=1, \pi_j(2)=2, \ldots, \pi_j(m)=m$. For $i=1,2,\ldots,m$, let

$$q_i = \Pr[Y_i = 1 \;\big|\; Y_1 = Y_2 = \cdots = Y_{i-1} = 0].$$

Note that q_i is not necessarily y_i/γ: if k and i lie in $N_{j'}$ for some clustered j', then Y_i and Y_k are negatively correlated. We will employ the following result of [31]:

$$\begin{aligned}\mathbf{E}[\sum_i c_{ij}X_{ij}] \;\;\leq\;\; & \left(\sum_{i=1}^{m} c_{ij} q_i \prod_{k<i}(1-q_k)\right) + \\ & \left(\prod_{k=1}^{m}(1-q_k)\right)\cdot 3R_j(\gamma). \qquad (3.16)\end{aligned}$$

Our plan is to bound this term altogether, rather than bounding the two summands separately as in [31].

Recall properties (A1)-(A5). Now, view v, $R_j(\gamma)$, and all the x^*_{ij}, x_{ij}, y_i etc. as fixed, with only the c_{ij} being variables subject to

$$\begin{aligned} 0 \;&\leq\; c_{1j} \leq \cdots \leq c_{vj} = R_j(\gamma); \\ R_j(\gamma) \;&\leq\; c_{(v+1)j} \leq \cdots \leq c_{mj}.\end{aligned}$$

Then, we have from (3.16) that $\mathbf{E}[\sum_i c_{ij}X_{ij}]/\overline{C}_j = \mathbf{E}[\sum_i c_{ij}X_{ij}]/\sum_{i=1}^m c_{ij}x^*_{ij}$ is bounded by a rational function h of the c_{ij} with the denominator being non-negative. Subject to the given constraints, it can be shown that this function h is maximized when there exists an u ($0 \leq u \leq m-1$) such that $c_{1j} = c_{2j} = \cdots = c_{uj} = 0$, and $c_{(u+1)j} = c_{(u+2)j} = \cdots = c_{mj} = R_j(\gamma)$. This simplifies our calculations and lets us complete the proof, using the definition of $R_j(\gamma)$ and the negative correlation among the random variables Y_i. We will present the complete proof of Lemma 3.2 in the full version of this work.

Setting $\gamma = 0.58$, we get (3.14) and (3.15). Next, we proceed as in Section 4.3 of [31] to use such a result for non-stochastic facility location, for two-stage stochastic facility location. Briefly, we start with the LP-relaxation of § 3.1. Recall the definition of $r^{(1)}_{A,j}$ from (3.5). All pairs (j,A) with $r^{(1)}_{A,j} \geq 1.73/(1.73+1.52)$ are declared selected in Stage I. For the Stage I decisions, we construct the facility-location instance $\mathcal{I}$ described immediately following (3.6), and solve it using our non-stochastic facility location algorithm above. (In contrast, the Stage-I problem is solved using the JMS algorithm in § 3.1.) In Stage II, we round separately for each scenario A as follows, just as in § 3.1. Construct a facility-location instance $\mathcal{I}_A$ with a unit-demand client for each $j \in A$ such that (j,A) was not selected in Stage I; each facility i has cost f^A_i. Now use the JMS algorithm as described in [17] to get an approximately optimal solution for $\mathcal{I}_A$. Just as in [31], it is easy to show that this new algorithm has a per-scenario expected approximation ratio of $1.73+1.52 = 3.25$.

Our discussion here is summarized by the following theorem:

THEOREM 3.2. *There is a randomized polynomial-time approximation algorithm $\mathcal{A}$ for two-stage stochastic facility location, with the following properties. $\mathcal{A}$ takes as input an instance $\mathcal{I}$ of two-stage stochastic facility location and any feasible solution $\mathcal{S}$ to the LP-relaxation of $\mathcal{I}$; $\mathcal{A}$ outputs a feasible solution for $\mathcal{I}$ in which the expected cost of any scenario A is at most 3.25 times the cost of A in $\mathcal{S}$.*

Acknowledgments. I thank Alok Baveja, David Shmoys, Chaitanya Swamy, and the SODA 2007 reviewers for their very helpful inputs. Thanks again to David and Chaitanya for writing the survey [33] that inspired this work. My gratitude also to Meena Mahajan and the Institute of Mathematical Sciences at Chennai, India, for their hospitality; part of this writing was completed at the IMSc.

References

[1] E. M. L. Beale. On minimizing a convex function subject to linear inequalities. *Journal of the Royal Statistical Society, Series B*, 17:173–184; discussion 194–203, 1955.

[2] M. Charikar, C. Chekuri, and M. Pál. Sampling bounds for stochastic optimization. *Proc. 9th RANDOM*, pages 257–269, 2005.

[3] F. Chudak and D. B. Shmoys. Improved approximation algorithms for the uncapacitated facility location problem. *SIAM Journal on Computing*, 33(1):1–25, 2003.

[4] G. B. Dantzig. Linear programming under uncertainty. *Mgmt. Sc.*, 1:197–206, 1955.

[5] B. Dean, M. Goemans, and J. Vondrak. Approximating the stochastic knapsack problem: the benefit of adaptivity. *Proc. 45th Annual IEEE Symposium on Foundations of Computer Science*, pages 208–217, 2004.

[6] B. Dean, M. Goemans, and J. Vondrak. Adaptivity and approximation for stochastic packing problems. In *Proc. Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 395–404, 2005.

[7] K. Dhamdhere, V. Goyal, R. Ravi, and M Singh. How to pay, come what may: approximation algorithms for demand-robust covering problems. *Proc. 46th Annual IEEE Symposium on Foundations of Computer Science*, pages 367–378, 2005.

[8] M. Dyer and L. Stougie. Computational complexity of stochastic programming problems. SPOR-Report 2005-11, Dept. of Mathematics and Computer Science, Eindhoven Technical University, Eindhoven, 2005.

[9] S. Guha and S. Khuller. Greedy strikes back: Improved facility location algorithms. *Journal of Algorithms*, 31(1):228–248, 1999.

[10] A. Gupta, M. Pál, R. Ravi, and A. Sinha. Boosted sampling: approximation algorithms for stochastic optimization. *Proc. 36th Annual ACM Symposium on Theory of Computing*, pages 417–426, 2004.

[11] A. Gupta, M. Pál, R. Ravi, and A. Sinha. What about Wednesday? Approximation algorithms for multistage stochastic optimization. *Proc. 8th APPROX*, pages 86–98, 2005.

[12] A. Gupta, R. Ravi, and A. Sinha. An edge in time saves nine: LP rounding approximation algorithms for stochastic network design. *Proc. 45th Annual IEEE Symposium on Foundations of Computer Science*, pages 218-227, 2004.

[13] N. Immorlica, D. Karger, M. Minkoff, and V. Mirrokni. On the costs and benefits of procrastination: approximation algorithms for stochastic combinatorial optimization problems. In *Proc. Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 684–693, 2004.

[14] K. Jain, M. Mahdian, E. Markakis, A. Saberi, and V. Vazirani. Greedy facility location algorithms analyzed using dual-fitting with factor-revealing LP. *Journal of the ACM*, 50(6):795–824, 2003.

[15] K. Jain, M. Mahdian, and A. Saberi. A new greedy approach for facility location problems. In *Proc. 34th Annual ACM Symposium on Theory of Computing*, pages 731–740, 2002.

[16] J. Linderoth, A. Shapiro, and R. K. Wright. The empirical behavior of sampling methods for stochastic programming. *Annals of Operations Research*, 142:215–241, 2006.

[17] M. Mahdian, Y. Ye, and J. Zhang. Approximation algorithms for metric facility location problems. *SIAM Journal on Computing*, 36:411–432, 2006.

[18] R. Möhring, A. Schulz, and M. Uetz. Approximation in stochastic scheduling: the power of LP based priority policies. *JACM*, 46:924–942, 1999.

[19] M. Naor and R. M. Roth. Optimal file sharing in distributed networks. *SIAM J. Comput.*, 24:158–183, 1995.

[20] A. Panconesi and A. Srinivasan. Randomized distributed edge coloring via an extension of the Chernoff-Hoeffding bounds. *SIAM J. Comput.*, 26:350–368, 1997.

[21] P. Raghavan and C. D. Thompson. Randomized rounding: a technique for provably good algorithms and algorithmic proofs. *Combinatorica*, 7:365–374, 1987.

[22] R. Ravi and A. Sinha. Hedging uncertainty: approximation algorithms for stochastic optimization problems. In *Proc. 10th IPCO*, pages 101–115, 2004.

[23] A. Ruszczynski and A. Shapiro. Editors, *Stochastic Programming*, Volume 10 of *Handbooks in Operations Research and Mgmt. Sc.*, North-Holland, Amsterdam, 2003.

[24] D. B. Shmoys and C. Swamy. An approximation scheme for stochastic linear programming and its application to stochastic integer programs. *JACM*, to appear. Preliminary version appeared as "Stochastic optimization is (almost) as easy as deterministic optimization" in *Proc. 45th Annual IEEE FOCS*, pages 228–237, 2004.

[25] D. B. Shmoys, É. Tardos, and K. I. Aardal. Approximation algorithms for facility location problems. In *Proc. 29th Annual ACM Symposium on Theory of Computing*, pages 265–274, 1997.

[26] A. M-C. So, J. Zhang, and Y. Ye. Stochastic combinatorial optimization with controllable risk aversion level. In *Proc. 9th APPROX*, pages 224–235, 2006.

[27] A. Srinivasan. Improved approximation guarantees for packing and covering integer programs. *SIAM J. Comput.*, 29:648–670, 1999.

[28] A. Srinivasan. New approaches to covering and packing problems. In *Proc. ACM-SIAM Symposium on Discrete Algorithms*, pages 567–576, 2001.

[29] M. Sviridenko. An improved approximation algorithm for the metric uncapacitated facility location problem. In *Proc. 9th IPCO*, pages 240–257, 2002.

[30] C. Swamy. *Approximation Algorithms for Clustering Problems.* Ph.D. thesis, Cornell University, Ithaca, NY, 2004.

[31] C. Swamy. *Approximation Algorithms for Clustering Problems.* Updated version of Ph.D. thesis, 2006. Available at *http://www.math.uwaterloo.ca/∼cswamy.*

[32] C. Swamy and D. B. Shmoys. Sampling-based approximation algorithms for multi-stage stochastic optimization. In *Proc. 46th Annual IEEE Symposium on Foundations of Computer Science*, pages 357–366, 2005.

[33] C. Swamy and D. B. Shmoys. Algorithms Column: Approximation Algorithms for 2-stage Stochastic Optimization Problems. *SIGACT News*, 37:33–46, 2006.

[34] B. Verweij, S. Ahmed, A. J. Kleywegt, G. L. Nemhauser, and A. Shapiro. The sample average approximation method applied to stochastic routing problems: a computational study. *Computational Optimization and Applications*, 24:289–333, 2003.