Submitted to Math Programming, manuscript No.

Improved Approximations for Two-stage Min-Cut and Shortest Path Problems under Uncertainty

Daniel Golovin · Vineet Goyal · Valentin Polishchuk · R. Ravi · Mikko Sysikaski

Received: date / Accepted: date

Abstract In this paper, we study the robust and stochastic versions of the two-stage min-cut and shortest path problems introduced in Dhamdhere et al. [6], and give approximation algorithms with improved approximation factors. Specifically, we give a 2-approximation for the robust min-cut problem and a 4-approximation for the stochastic version. For the two-stage shortest path problem, we give a 3.39-approximation for the robust version and 6.78-approximation for the stochastic version. Our results significantly improve the previous best approximation factors for the problems. In particular, we provide the first constant-factor approximation for the stochastic min-cut problem.

Our algorithms are based on guess and prune strategy that crucially exploits the nature of the robust and stochastic objective. In particular, we guess the worst-case second stage cost and based on the guess, select a subset of *costly* scenarios for the first-stage solution to address. The second-stage solution for any scenario is simply the min-cut (or shortest

Part of this work was done while D. Golovin was affiliated with Carnegie Mellon University and was supported by NSF grants CCR-0122581 and IIS-0121678. V. Goyal is supported by NSF grant CMMI-120116. R. Ravi is supported by NSF grants CCF-1218382 and CCF-1143998. V. Polishchuk is supported by the Academy of Finland grant 138520. M. Sysikaski is supported by research funds of University of Helsinki.

D. Golovin
Google, Inc.
E-mail: golovin@gmail.com

V. Goyal
Department of Industrial Engineering and Operations Research
Columbia University
E-mail: vgoyal@ieor.columbia.edu

V. Polishchuk
University of Helsinki
E-mail: valentin.polishchuk@helsinki.fi

R. Ravi
Tepper School of Business
Carnegie Mellon University
E-mail: ravi@andrew.cmu.edu

M. Sysikaski
University of Helsinki
E-mail: mikko.sysikaski@helsinki.fi

path) problem in the residual graph. The key contribution is to show that there is a near-optimal first stage solution that completely satisfies the subset of costly scenarios that are selected by our procedure. While the guess and prune strategy is not directly applicable for the stochastic versions, we show that using a novel LP formulation, we can adapt a guess and prune algorithm for the stochastic versions. Our algorithms based on the guess and prune strategy provide insights about the applicability of this approach for more general robust and stochastic versions of combinatorial problems.

1 Introduction

In this paper, we study the two-stage robust and stochastic versions of two classical combinatorial optimization problems, namely, minimum-cut and shortest path under demand uncertainty, where cut or connectivity requirements are uncertain in addition to the objective coefficient uncertainty. The goal is to build a two-stage solution to minimize the worst-case or the expected cost, where in the first-stage the cost is lower but the constraints are uncertain, and in the second-stage the constraints are known but the cost is higher. This tradeoff between cost and uncertainty is common in most decision problems in the real world where parameters are often uncertain in the optimization phase. Both stochastic and robust optimization have been extensively studied in the literature to model decision problems under uncertainty. We refer the reader to several textbooks including Infanger [15], Kall and Wallace [16], Prekopa [18], Shapiro [20], Shapiro et al. [21] and the references therein for a comprehensive review of stochastic optimization, and to the survey by Bertsimas et al. [3] and the book by Ben-Tal et al. [2] and the references therein for an extensive review of the literature in robust optimization.

Stochastic optimization models for general covering problems were introduced and studied in several recent papers (see Gupta et al. [13], Immorlica et al. [14], Ravi and Sinha [19], Shmoys and Swamy [22]). Dhamdhere et al. [6] introduce the demand-robust version of two-stage covering problems where the objective is to minimize the worst-case cost instead of the expected cost as in the stochastic version. They prove a structural result about the first-stage solution in this model and use it to give approximation algorithms for several two-stage covering problems including set cover, multi-cut, Steiner tree and facility location. In this paper, we consider the robust and stochastic versions of min-cut and shortest path, and give improved approximation algorithms. These special cases of the multi-cut and Steiner tree problems allow us to focus on the complexity of robust and stochastic modeling. Both min-cut and shortest path problems are polynomial time solvable in the deterministic case but are APX-hard in the two-stage model (see Khandekar et al. [17] and Gupta et al. [13]). We propose algorithms with significantly better approximation ratios as compared to the algorithms in Dhamdhere et al. [6] and Gupta et al. [13]. More importantly, our algorithms are based on a guess and prune strategy that exploit crucial properties of the robust and stochastic objective, thus providing interesting insights for developing algorithms for more general problems in these models.

1.1 Model and Notations

Let us formally introduce the two-stage min-cut and shortest path problems.

Two-stage Minimum cut problems. We consider both robust and stochastic versions of the two-stage min-cut problem under demand uncertainty. We are given an undirected graph

$G = (V, E)$ with edge costs $c_f : E \to \mathbb{R}_+$ and a root vertex r. Unlike the deterministic min-cut problem, the terminal that needs to be separated from r is uncertain and is given by a list of k scenarios, only one of which is realized in the second stage. Each scenario $i = 1, \ldots, k$ is specified by the terminal t_i and an inflation factor σ_i which is the factor by which each edge becomes costlier in the second stage in scenario i. The goal is to build a two-stage solution, i.e., select a set of edges E_f in the first-stage and, for each scenario i, a set of recourse edges E_s^i in the second-stage such that $E_f \cup E_s^i$ separate r from t_i for all scenarios $i = 1, \ldots, k$. Note the tradeoff between cost and uncertainty: in the first-stage, the edge costs are small but the terminal that needs to be separated from r is uncertain. In the second-stage, we know the terminal but the edge costs are higher by the corresponding inflation factor. Let $[k]$ denote $\{1, \ldots, k\}$.

Robust Min-Cut Problem, (Π_{RMC}). In the robust min-cut problem, the goal is to find a two-stage min-cut solution such that the worst-case cost over all scenarios is minimized, i.e.,

$$\begin{aligned} \min_{E_f, E_s^i \subseteq E, i \in [k]} \; & c_f(E_f) + \max_{i=1}^{k} \; \sigma_i \cdot c_f(E_s^i) \\ \text{s.t. } & E_f \cup E_s^i \text{ separate } r \text{ from } t_i. \end{aligned} \tag{1.1}$$

Stochastic Min-Cut Problem, (Π_{SMC}). In the stochastic min-cut problem, for each scenario $i \in [k]$, we are also given its probability of occurrence, p_i such that $p_1 + \ldots + p_k = 1$. The goal is to find a two-stage min-cut solution such that the total expected cost is minimized, i.e.,

$$\begin{aligned} \min_{E_f, E_s^i \subseteq E, i \in [k]} \; & c_f(E_f) + \sum_{i=1}^{k} \; p_i \sigma_i \cdot c_f(E_s^i) \\ \text{s.t. } & E_f \cup E_s^i \text{ separate } r \text{ from } t_i. \end{aligned} \tag{1.2}$$

Two-stage shortest path problems. Similar to the robust and stochastic min-cut problems, we are given an undirected graph $G = (V, E)$ with edge costs $c_f : E \to \mathbb{R}_+$, a root vertex r, and a list of k scenarios, only one of which realizes in the second-stage. Each scenario $i = 1, \ldots, k$ specifies the terminal t_i that must be connected to r if that scenario realizes, and an inflation factor σ_i for edge costs. The goal is to build a two-stage solution, i.e., select a set of edges E_f in the first-stage and for each scenario i, a set of recourse edges E_s^i in the second-stage such that $E_f \cup E_s^i$ connect r and t_i for all scenarios $i = 1, \ldots, k$.

Robust shortest path problem, (Π_{RSP}). In the robust shortest path problem, the goal is to find a two-stage shortest path solution such that the worst-case cost over all scenarios is minimized, i.e.,

$$\begin{aligned} \min_{E_f, E_s^i \subseteq E, i \in [k]} \; & c_f(E_f) + \max_{i=1}^{k} \; \sigma_i \cdot c_f(E_s^i) \\ \text{s.t. } & E_f \cup E_s^i \text{ connect } r \text{ and } t_i. \end{aligned} \tag{1.3}$$

Stochastic shortest path problem, (Π_{SSP}). In the stochastic shortest path problem, we are also given probabilities of occurrence, $p_1, \ldots, p_k$ for all scenarios, such that $p_1 + \ldots + p_k = 1$. The goal is to find a two-stage solution such that the total expected cost is minimized, i.e.,

$$\begin{aligned} \min_{E_f, E_s^i \subseteq E, i \in [k]} \; & c_f(E_f) + \sum_{i=1}^{k} \; p_i \sigma_i \cdot c_f(E_s^i) \\ \text{s.t. } & E_f \cup E_s^i \text{ connect } r \text{ and } t_i. \end{aligned} \tag{1.4}$$

1.2 Our Contributions

In this paper, we present improved approximation algorithms for the robust and stochastic versions of the min-cut and shortest path problems introduced above. Our main contributions are the following.

1. We present a 2-approximation for the robust min-cut problem which improves on the $O(\log n)$-approximation for the problem in [6], and also on the $(1+\sqrt{2})$-approximation that appears in a conference version of this paper [10]. Khandekar et al. [17] show that the robust min-cut problem is APX-hard even for the case of three scenarios and uniform inflation factor.

   Our algorithm is based on a guess and prune strategy, where we guess the worst-case second-stage cost of an optimal solution. Based on the guess, we identify a set of *costly* scenarios (corresponding to terminals whose min-cut from r has a high cost), and we compute the first-stage solution that separates r from these terminals of costly scenarios. The second-stage solution for any scenario i is simply an r-t_i min-cut in the residual graph. The analysis is based on the structural properties of the minimum cuts. In particular, we use a novel charging argument using Gomory–Hu tree representation of pairwise min-cuts [11].

2. For the stochastic min-cut, we present a 4-approximation. Our algorithm is the first constant-factor approximation for the stochastic min-cut problem. Since the expected objective for the stochastic problem depends on the second-stage costs for all the scenarios and not just the worst-case cost, a guess and prune strategy is not directly applicable. However, we propose a novel approximate 0-1 formulation for the stochastic min-cut which can be rounded to give a 4-approximation for the problem.

3. We also give a $(\gamma+2)$-approximation for the robust shortest path problem, where γ is the best approximation factor for the Steiner tree problem on undirected graphs. The best $\gamma = 1.39$ is due to Byrka et al. [4] which implies a 3.39-approximation for the problem. This significantly improves on the 30-approximation for the problem in [6] and 7.1-approximation in the preliminary version of this paper [10]. The robust (and stochastic) shortest path problem is NP-hard as it contains the Steiner tree problem (if inflation factors = ∞, then the optimal solution is an optimal Steiner tree over all scenario terminals).

   Similarly to the min-cut problem, the algorithm is based on a guess and prune strategy where we guess the worst-case second-stage cost of an optimal solution. However, the selection of *costly* scenarios to build the first-stage solution is quite different. The cost of building a feasible solution for all the costly scenarios in the first-stage can be potentially high. Therefore, we need to carefully select only a subset of costly scenarios for the first stage solution to achieve the desired approximation factor.

4. For the stochastic shortest path problem, we give a $2(\gamma+2)$-approximation which implies a 6.78-approximation (since $\gamma = 1.39$ due to Byrka et al. [4]). This improves on the 30-approximation for the problem in [13]. Similarly to the case of stochastic min-cut, we use an LP formulation for the problem to adapt to a guess and prune strategy. However, we lose an approximation factor of 2 as compared to the robust version (similar to the case of min-cut problem).

As mentioned above, in all the algorithms, we use a guess and prune strategy in some form: we either directly guess the worst-case cost (for the robust versions) or use an LP formulation (for the stochastic versions) to find a set of *costly* scenarios, and then build a feasible solution for these scenarios in the first stage. We exploit the form of robust and stochastic objectives crucially and our analysis can provide insights for using this technique for more general robust and stochastic combinatorial problems.

1.3 Related Work

Stochastic and robust versions of combinatorial problems have been studied extensively in the literature. Two-stage stochastic versions of the combinatorial problems were introduced in Ravi and Sinha [19] and Immorlica et al. [14] who gave approximation algorithms for several covering and packing problems. This led to a series of papers in this area including Gupta et al. [13] who give sampling based algorithms for polynomial number of scenarios, Shmoys and Swamy [22] who consider the case of exponential number of scenarios, and Charikar et al. [5] who propose a sample average approximation for two-stage stochastic combinatorial problems.

Dhamdhere et al. [6] introduced the robust version of two-stage combinatorial covering problems under uncertainty and gave approximation algorithms for several covering problems including set cover, min-cut, multi-cut, Steiner tree, and facility location. In particular, they give a 30-approximation for the robust Steiner tree problem, $O(\log n)$-approximation for the robust min-cut problem, and $O(\log n \log\log n)$-approximation for the robust multi-cut problem. The techniques in that paper also give first approximation algorithms for the stochastic min-cut and stochastic multi-cut problems with the same approximation factors as the robust versions. In a conference version of this paper [10], we present improved approximations for the robust min-cut and robust shortest path problems: $(1+\sqrt{2})$-approximation for the robust min-cut problem and 7.1-approximation for the robust shortest path problem. Feige et al. [7] and later Khandekar et al. [17] consider the case of exponential scenarios that are specified implicitly (all subsets of at most p terminals is a possible scenario) and give approximations for set cover, Steiner tree and facility location problems. Gupta et al. [12] generalize the result for exponential scenarios and give approximation algorithms for covering problems based on the guess and prune idea.

Outline. The rest of the paper is organized as follows. In Section 2, we present the 2-approximation for the robust min-cut problem and in Section 3, we describe our 4-approximation for the stochastic min-cut problem. In Sections 4 and 5, we describe our approximations for the robust and stochastic shortest path problems respectively.

2 Two-stage Robust Min-Cut

In this section, we present our algorithm for the two-stage robust min-cut problem, Π_{RMC} defined in (1.1). We present a 2-approximation for this problem improving the previous best approximation factor of $(1+\sqrt{2})$ in [10]. To motivate our approach, let us consider the special case of the robust min-cut problem when the graph is a tree. Suppose we know the maximum second-stage cost for an optimal solution is C. Any terminal $t_i,\ i = 1,\ldots,k$ whose min-cut from r costs more than $\frac{C}{\sigma_i}$ (with respect to the first stage costs) should be separated from r in the first stage by this optimal solution. Otherwise, the second-stage cost for scenario i will be more than C. Furthermore, if the min-cut between r and $t_i, i \in [k]$ is less than or equal to $\frac{C}{\sigma_i}$, this scenario can be handled independently in the second-stage within a cost at most C. Therefore, we can ignore scenario i in constructing the first-stage solution. This implies that if we know the value of C, we can identify exactly the set of terminals U that should be disconnected from r in the first stage. While we do not know the value of C, there are only a small number of values that are critical. Let $\text{cut-cost}(t_i)$ denote the cost of the minimum cut between r and t_i in G with respect to first-stage cost function c_f. Notice that for a particular value of C,

$$U = \{t_i \mid \sigma_i \cdot \text{cut-cost}(t_i) > C, i = 1, \ldots, k\}.$$

Assume wlog. that the scenarios are ordered such that

$$\sigma_1 \cdot \text{cut-cost}(t_1) \geq \sigma_2 \cdot \text{cut-cost}(t_2) \geq \ldots \geq \sigma_k \cdot \text{cut-cost}(t_k).$$

Therefore, for any value of C, $U = \emptyset$ or $U = \{t_1, \ldots, t_j\}$ for some $j = 1, \ldots, k$ which implies that we need to try only $k+1$ values of C to find the best solution. This motivates the following algorithm described in Figure 1.

Algorithm $\mathcal{A}_{\mathsf{RMC}}$ for the Robust Min-Cut Problem

Input: Undirected graph $G = (V, E)$, root r, terminals $T = \{t_1, t_2, \ldots, t_k\}$, inflation factors $\sigma_1, \ldots, \sigma_k$.

1. Reorder terminals such that

$$\sigma_1 \cdot \text{cut-cost}(t_1) \geq \sigma_2 \cdot \text{cut-cost}(t_2) \geq \ldots \geq \sigma_k \cdot \text{cut-cost}(t_k).$$

2. Initialize $z \leftarrow \infty$.
3. For $j = 0, \ldots, k$
   (a) Let $U = \{t_1, \ldots, t_j\}$. (Note $U = \emptyset$ when $j = 0$.)
   (b) Compute the first stage solution:

$$\hat{E}_f \leftarrow \text{ minimum } r\text{-}U \text{ cut in } G.$$

   (c) Compute the second stage solution for scenario i:

$$\hat{E}_s^i \leftarrow \text{ minimum } r\text{-}t_i \text{ cut in } G \setminus \hat{E}_f.$$

   (d) If $z > c_f(\hat{E}_f) + \max_{i=1}^k \sigma_t \cdot c_f(\hat{E}_s^i)$

$$\begin{aligned} z &\leftarrow c_f(\hat{E}_f) + \max_{i=1}^{k} \sigma_t \cdot c_f(\hat{E}_s^i) \\ E_f &\leftarrow \hat{E}_f \\ E_s^i &\leftarrow \hat{E}_s^i \text{ for } i = 1, \ldots, k. \end{aligned}$$

4. **Return:** E_f, E_s^i for $i = 1, \ldots, k$.

Fig. 1 Algorithm for the Robust Min-Cut Problem. It runs in $\tilde{O}(k^2 mn)$ time on undirected graphs using the max flow algorithm of Goldberg and Tarjan [9] to find minimum cuts.

2.1 Special Case: G is a Tree

As a warmup, let us first prove that Algorithm $\mathcal{A}_{\mathsf{RMC}}$ solves the robust min-cut problem optimally when the underlying graph is a tree.

Theorem 2.1 *Algorithm $\mathcal{A}_{\mathsf{RMC}}$ computes an optimal solution for the robust min-cut problem when G is a tree.*

Proof Consider an optimal solution $E_f^*, E_s^{i*}, i = 1, \ldots, k$ for the given problem instance. The optimal objective value is

$$z^* = c_f(E_f^*) + \max_{i=1}^{k} \ \sigma_i \cdot c_f(E_s^{i*}). \tag{2.1}$$

Note that for terminal $t_i, i = 1, \ldots, k$ that is not separated from r in $G \setminus E_f^*$, E_s^{i*} is the minimum-cost edge on the unique path between r and t_i in $G \setminus E_f^*$ (which is also the unique path between r and t_i in G). Therefore,

$$c_f(E_s^{i*}) = \begin{cases} \text{cut-cost}(t_i), & \text{if } t_i \text{ is connected to } r \text{ in } G \setminus E_f^* \\ 0, & \text{otherwise.} \end{cases} \tag{2.2}$$

Let j_1 be the smallest index j such that t_j is not separated from r in $G \setminus E_f^*$. Let

$$U = \{t_1, \ldots, t_{j_1-1}\},$$

and let

$$\hat{E}_f \leftarrow \text{minimum } r\text{-}U \text{ cut in } G$$

$$\hat{E}_s^i \leftarrow \begin{cases} \text{minimum } r\text{-}t_i \text{ cut in } G, & i = j_1, \ldots, k \\ \emptyset, & \text{otherwise.} \end{cases}$$

Note that E_f^* separates all terminals in U from r. Therefore,

$$c_f(E_f^*) \geq c_f(\hat{E}_f), \tag{2.3}$$

since $\hat{E}_f$ is a minimum cut separating U from r. Also,

$$\max_{i=1}^{k} \ \sigma_i \cdot c_f(E_s^{i*}) \geq \sigma_{j_1} \cdot c_f(E_s^{j_1}) = \sigma_{j_1} \cdot \text{cut-cost}(t_{j_1}), \tag{2.4}$$

where the second equality follows from (2.2). Furthermore,

$$\begin{aligned} \max_{i=1}^{k} \sigma_i \cdot c_f(\hat{E}_s^i) &= \max_{i=j_1}^{k} \ \sigma_i \cdot c_f(\hat{E}_s^i) \\ &= \max_{i=j_1}^{k} \ \sigma_i \cdot \text{cut-cost}(t_i) \\ &= \sigma_{j_1} \cdot \text{cut-cost}(t_{j_1}), \end{aligned}$$

where the first equality follows as $\hat{E}_s^i = \emptyset$ for all $i < j_1$ and the last equality follows as the terminals are ordered in decreasing $\sigma_i \cdot \text{cut-cost}(t_i)$. Therefore, we can compare the cost of the solution $\hat{E}_f, \hat{E}_s^i, i = 1, \ldots, k$ as follows.

$$\begin{aligned} c_f(\hat{E}_f) + \max_{i=1}^{k} \ \sigma_i \cdot c_f(E_s^i) &= c_f(\hat{E}_f) + \sigma_{j_1} \cdot \text{cut-cost}(t_{j_1}) \\ &\leq c_f(E_f^*) + \max_{i=1}^{k} \ \sigma_i \cdot c_f(E_s^{i*}) \\ &= z^*, \end{aligned}$$

where the second inequality follows from (2.3) and (2.4) and the last equation follows from (2.1). Note that Algorithm $\mathcal{A}_{\mathsf{RMC}}$ will examine the choice of $U = \{t_1, \ldots, t_{j_1 - 1}\}$ and therefore, computes an optimal solution of the robust min-cut problem on trees. $\square$

2.2 General Graphs

We show that, surprisingly, Algorithm $\mathcal{A}_{\mathsf{RMC}}$ in Figure 1 performs well even for the case of general graphs. Algorithm $\mathcal{A}_{\mathsf{RMC}}$ uses the following paradigm: any scenario is satisfied completely in the first stage or is budgeted to be satisfied completely in the second stage (without any help from the first-stage solution). While this leads to an optimal algorithm for trees (Theorem 2.1), it is not the case for general graphs. If for some terminal t_i the minimum r-t_i cut has a cost more than $\frac{C}{\sigma_i}$ (where C is the maximum second-stage cost for an optimal solution), then we can only infer that the optimal solution should "help" this terminal in the first stage to keep the second-stage cost at most C, unlike in trees where the first-stage solution must have separated t_i from r. However, we show that Algorithm $\mathcal{A}_{\mathsf{RMC}}$ is a good approximation even for the case of general graphs. The analysis is based on a novel charging scheme that exploits the structure of minimum cuts. Let us introduce a few definitions and notations.

Let $E_f^*, E_s^{i*},\ i = 1, \ldots, k$ be an optimal two-stage solution (denoted by OPT). The optimal objective value is

$$z^* = c_f(E_f^*) + \max_{i=1}^{k}\ \sigma_i \cdot c_f(E_s^{i*}) = c_f(E_f^*) + C,$$

where C is the worst-case second-stage cost of OPT.

Definition 1 A scenario i is referred to as *costly* if the r-t_i min-cut cost is more than $\frac{2C}{\sigma_i}$.

Let S denote the set of terminals corresponding to costly scenarios, i.e.,

$$S = \{t_i \mid \sigma_i \cdot \text{cut-cost}(t_i) > 2C\}. \tag{2.5}$$

Note that $S = \{t_1, \ldots, t_j\}$ for some $j = 0, \ldots, k$ since the terminals are sorted in decreasing order of $\sigma_i \cdot \text{cut-cost}(t_i)$. In any optimal solution, a costly scenario that is not satisfied completely in the first stage needs some help from the first-stage solution to keep the worst-case second-stage cost to at most C.

For any two vertices $u, v \in V$, let

$$\text{cut-cost}_G(u, v) := \text{ cost of } u\text{-}v \text{ min-cut in graph } G.$$

For any $T \subset V$, let

$$\delta_G(T) = \{e = (u, v) \in E(G) \mid u \in T, v \notin T\}$$

denote the edges that cut out T. Finally, let $G' = (V, E \setminus E_f^*)$ be the graph obtained by removing E_f^* from G.

Gomory–Hu Tree. We use the Gomory–Hu tree representation of pairwise min-cuts in an undirected graph to exploit the laminar structure of minimum cuts. For graph $G = (V, E)$, a Gomory–Hu tree is a tree $H = (V, F)$ on vertices V with edges F and a cost function $c_h : F \to \mathbb{R}_+$ that satisfies the following properties.

1. For any two vertices $u, v \in V$, let $\mathcal{P}(u,v)$ denote the unique path from u to v in T. Then, the cost of the minimum u-v cut in graph G is the cost (with respect to c_h) of the minimum cost edge on $\mathcal{P}(u,v)$, i.e.,

$$\text{cut-cost}_G(u,v) = \min_{e \in \mathcal{P}(u,v)} c_h(e).$$

2. If

$$e_{uv} = \text{argmin}_{e \in \mathcal{P}(u,v)} c_h(e),$$

then the two connected components obtained by removing e_{uv} from T form a u-v min-cut in G.

Approximate Solution. We construct the approximate first-stage and second-stage solution, $E_f, E_s^i,\ i = 1, \dots, k$ for the robust min-cut problem on G as follows.

$$\begin{aligned} E_f &\leftarrow \text{ minimum } r\text{-}S \text{ cut in } G \\ E_s^i &\leftarrow \text{ minimum } r\text{-}t_i \text{ cut in } G \setminus E_f. \end{aligned} \tag{2.6}$$

Note that since $S = \{t_1, \dots, t_j\}$ for some $j = 0, \dots, k$, Algorithm $\mathcal{A}_{\mathsf{RMC}}$ considers this solution in one of the iterations. We show that the total worst-case cost of the above solution is at most $2z^*$. First, we show that the worst-case second-stage cost is at most $2C$.

Lemma 2.2 *Let E_f, E_s^i, for all $i = 1, \dots, k$ be as defined in* (2.6). *Then for all $i = 1, \dots, k$, the second-stage cost for scenario i at most $2C$, i.e.,*

$$\sigma_i \cdot c_f(E_s^i) \le 2C,\ i = 1, \dots, k.$$

Proof If $t_i \in S$, then $E_s^i = \emptyset$ since E_f separates r from t_i, and $c_f(E_s^i) = 0 \le 2C$. If $t_i \notin S$, then

$$\sigma_i \cdot c_f(E_s^i) = \sigma_i \cdot c_f(\text{ minimum } r\text{-}t_i \text{ cut in } G \setminus E_f) \le \sigma_i \cdot \text{cut-cost}(t_i),$$

which is at most $2C$ from the definition of S (2.5). □

We next show that the cost of the first-stage solution E_f defined in (2.6) is at most twice the cost of E_f^*. We prove this by constructing a r-S cut whose cost is at most $2c_f(E_f^*)$. The cost analysis relies on the laminar structure of minimum cuts and we use a Gomory–Hu tree to represent the pairwise min-cuts in $G' = G \setminus E_f^*$. Let $H = (V, F)$ be the Gomory–Hu tree for G' with respect to edge costs c_f, and let $c_h : F \to \mathbb{R}_+$ be the cost function on edges F. For any terminal $t \in S$, let

$$\begin{aligned} e_t &\in \text{argmin}_{e \in \mathcal{P}(t,r)}\ c_h(e) \\ U_t &\leftarrow \text{ Component containing } t \text{ in } H \setminus \{e_t\} \end{aligned} \tag{2.7}$$

where for any vertex $v \in V$, $\mathcal{P}(r,v)$ denotes the unique path from r to v in H. Recall that $c_h(e_t)$ is the cost of the minimum r-t cut in $G' = G \setminus E_f^*$, and that the two components obtained by removing e_t from H define an r-t min-cut.

We first prove the following lemma.

Lemma 2.3 *For any $t \in S$,*

$$c_f(\delta_G(U_t) \setminus E_f^*) = c_h(e_t) \le c_f(\delta_G(U_t) \cap E_f^*),$$

where e_t and U_t are as defined in (2.7).

Proof Let σ_t denote the inflation factor of the scenario corresponding to terminal $t \in S$. Note that $c_h(e_t)$ is the cost (with respect to c_f) of the second-stage solution in OPT for scenario corresponding to terminal t. Therefore,

$$c_f(\delta_G(U_t) \setminus E_f^*) = c_h(e_t) \leq \frac{C}{\sigma_t}, \tag{2.8}$$

as the second stage cost of OPT is at most C.

Since $\delta_G(U_t)$ is an r-t cut, and $t \in S$,

$$c_f(\delta_G(U_t)) \geq \text{cut-cost}(t) > \frac{2C}{\sigma_t}. \tag{2.9}$$

Now,

$$\delta_G(U_t) = \left(\delta_G(U_t) \cap E_f^*\right) \uplus (\delta_G(U_t) \setminus E_f^*),$$

which implies

$$c_f(\delta_G(U_t) \cap E_f^*) = c_f(\delta_G(U_t)) - c_f(\delta_G(U_t) \setminus E_f^*) > \frac{2C}{\sigma_t} - \frac{C}{\sigma_t} = \frac{C}{\sigma_t},$$

where the inequality follows from (2.8), (2.9). Therefore,

$$c_f(\delta_G(U_t) \setminus E_f^*) \leq \frac{C}{\sigma_t} < c_f(E_f^* \cap \delta_G(U_t)).$$

□

Now, we are ready to prove the main theorem.

Theorem 2.4 *Algorithm $\mathcal{A}_{\mathsf{RMC}}$ gives a 2-approximation for the two-stage robust min-cut problem, Π_{RMC}.*

Proof We will construct a low-cost r-S cut in G. In particular, consider the following cut,

$$\hat{E}_f = \delta_G\left(\bigcup_{t \in S} U_t\right).$$

Clearly, $\hat{E}_f$ separates r from S since $t \in U_t$ for every $t \in S$. Note that $U_t, t \in S$ form a laminar family of sets, i.e., for any two sets U_{t_1}, U_{t_2} where $t_1, t_2 \in S$, either one is contained in the other (i.e. $U_{t_1} \subseteq U_{t_2}$ or $U_{t_2} \subseteq U_{t_1}$) or they are disjoint (i.e. $U_{t_1} \cap U_{t_2} = \emptyset$). Therefore, we can find a minimal subset of terminals $\mathcal{T} \subseteq S$ such that

$$\bigcup_{t \in \mathcal{T}} U_t = \bigcup_{t \in S} U_t, \text{ and } U_t \cap U_{t'} = \emptyset, \ \forall t, t' \in \mathcal{T}.$$

We have

$$\hat{E}_f = \delta_G\left(\bigcup_{t \in \mathcal{T}} U_t\right) = \left(\bigcup_{t \in \mathcal{T}} \delta_G(U_t)\right) \setminus E_G\left[\bigcup_{t \in \mathcal{T}} U_t\right], \tag{2.10}$$

where for any $V' \subset V$, $E_G[V']$ denotes the set of edges in G induced between vertices in V'. We now show that $c_f(\hat{E}_f) \leq 2c_f(E_f^*)$, where E_f^* is an optimal first-stage solution. Consider any $e \in E_f^*$. Since $U_t, t \in \mathcal{T}$ are disjoint, $e \in \delta_G(U_t)$ for at most two terminals in $\mathcal{T}$. Let

$$\begin{aligned} E_{f,1}^* &= \{e \in E_f^* |\ e \in \delta_G(U_t) \text{ for exactly one terminal } t \in \mathcal{T}\}, \\ E_{f,2}^* &= \{e \in E_f^* |\ e \in \delta_G(U_t) \text{ for exactly two terminals } t, t' \in \mathcal{T}\}. \end{aligned} \tag{2.11}$$

Note that $E^*_{f,2} \subseteq E_G[\cup_{t\in\mathcal{T}} U_t]$ since any $e \in E^*_{f,2}$ is contained in $\delta_G(U_t)$ and $\delta_G(U_{t'})$ for distinct $t, t' \in \mathcal{T}$. Combining with (2.10), we have

$$\hat{E}_f \subseteq \left(\bigcup_{t\in\mathcal{T}} \delta_G(U_t)\right) \setminus E^*_{f,2} = \left(\bigcup_{t\in\mathcal{T}} \delta_G(U_t) \setminus E^*_f\right) \cup E^*_{f,1}. \tag{2.12}$$

Therefore,

$$\begin{aligned}
c_f(\hat{E}_f) &\leq \left(\sum_{t\in\mathcal{T}} c_f(\delta_G(U_t) \setminus E^*_f)\right) + c_f(E^*_{f,1}) \\
&\leq \left(\sum_{t\in\mathcal{T}} c_f(\delta_G(U_t) \cap E^*_f)\right) + c_f(E^*_{f,1}) && (2.13) \\
&= \sum_{t\in\mathcal{T}} \big(c_f(\delta_G(U_t) \cap E^*_{f,1}) + c_f(\delta_G(U_t) \cap E^*_{f,2})\big) + c_f(E^*_{f,1}) && (2.14) \\
&\leq 2c_f(E^*_{f,1}) + 2c_f(E^*_{f,2}) && (2.15) \\
&\leq 2c_f(E^*_f),
\end{aligned}$$

where the first inequality follows from (2.12). Inequality (2.13) follows from Lemma 2.3. Equality (2.14) follows since $E^*_{f,1}$ and $E^*_{f,2}$ are disjoint and any $e \in \delta_G(U_t \cap E^*_f)$ belongs to either $E^*_{f,1}$ or $E^*_{f,2}$. Inequality (2.15) follows since

$$\sum_{t\in\mathcal{T}} c_f(\delta_G(U_t) \cap E^*_{f,1}) = c_f(E^*_{f,1}), \; \sum_{t\in\mathcal{T}} c_f(\delta_G(U_t) \cap E^*_{f,2}) = 2c_f(E^*_{f,2}).$$

This completes the proof. □

Algorithm $\mathcal{A}_{\mathsf{RMC}}$ computes $O(k^2)$ minimum cuts and therefore, the running time is $\tilde{O}(k^2 mn)$ time on undirected graphs using the max flow algorithm of Goldberg and Tarjan [9] to find minimum cuts.

3 Two-stage Stochastic Min-Cut Problem

In this section, we present a constant-factor approximation for the two-stage stochastic min-cut problem, Π_{SMC} defined in (1.2). Here, the goal is to minimize the total expected cost of the two-stage solution instead of the worst-case cost over all scenarios as in the robust version. We present a 4-approximation for this problem. This is the first constant-factor approximation for the problem improving upon an $O(\log n)$-approximation given in [6]. The stochastic min-cut problem differs significantly from the robust version. In an optimal solution $E_f, E^i_s, i = 1, \ldots, k$, for the robust version of the problem, the second-stage solution E^i_s need not be an optimal r-t_i cut in $G \setminus E_f$, especially if scenario i is not one of the worst-case scenarios. Therefore, we can assume the second-stage cost for all scenarios is identical and we just need to ensure that there is a feasible second-stage solution for all scenarios with cost at most the worst-case cost. This property allows to identify the set of terminals that should be separated from the root in an approximate first stage solution by guessing just the worst-case second-stage cost in an optimal solution. However, this property is not true for the stochastic version. In an optimal solution $E^*_f, E^{i*}_s, i = 1, \ldots, k$ for the stochastic version, the second-stage solution, E^{i*}_s is a r-t_i min-cut in $G \setminus E^*_f$ for all scenarios i. Therefore, we can not use a guess and prune algorithm (with small number of guesses) to identify a set of terminals to disconnect in an approximate first-stage solution. However, we give a novel LP formulation to prune the scenarios and construct an approximate solution.

Let us first introduce an approximate 0-1 formulation for the problem. For any $i = 1, \ldots, k$, let $y_i \in \{0,1\}$ denote whether terminal t_i is separated from r in the first stage or not. For each edge $e \in E$, let $x_e \in \{0,1\}$ denote whether edge e is selected in the first stage solution or not. Let $\mathcal{P}_G(u,v)$ denote the set of paths from u to v in graph G. Also, as earlier, let $\text{cut-cost}_G(r,t_i)$ denote the cost of the minimum r,t_i-cut in G with respect to the cost function c_f. Consider the following integer program.

$$\begin{aligned} z_{IP} = \min \ & \sum_{e \in E} c_f(e) \cdot x_e + \sum_{i=1}^{k} \sigma_i p_i \cdot \text{cut-cost}_G(r,t_i) \cdot (1 - y_i) \\ \text{s.t.} \ & \sum_{e \in P} x_e \geq y_i, \ \forall P \in \mathcal{P}_G(r,t_i), \forall i = 1, \ldots, k \\ & x_e \in \{0,1\}, \ \forall e \in E \\ & y_i \in \{0,1\}, \ \forall i = 1, \ldots, k. \end{aligned} \tag{3.1}$$

Note that the above integer program is not an exact formulation for the stochastic min-cut problem. If a terminal $t_i, i = 1, \ldots, k$ is not separated from r by the first-stage solution x_e, i.e., $y_i = 0$, then we incur the cost of r-t_i min-cut in G in the second-stage instead of the min-cut in the modified graph after removing the first-stage edges. However, we show that this formulation is a good approximation of the stochastic min-cut problem. In particular, we show that z_{IP} is at most twice the optimal objective of the stochastic min-cut problem in the following lemma.

Lemma 3.1 *Let $E_f^*, E_s^{i*}, i = 1, \ldots, k$ be an optimal solution for the stochastic min-cut problem and z_{IP} be as defined in* (3.1). *Then*

$$z_{IP} \leq 2 \cdot \left(c_f(E_f^*) + \sum_{i=1}^{k} \sigma_i p_i \cdot c_f(E_s^i) \right) = 2z^*$$

Proof We will construct a feasible solution for (3.1) from the optimal stochastic solution with cost at most $2z^*$. As before, we use a Gomory–Hu tree representation of the pairwise min-cuts. In particular, let $H = (V, F)$ be the Gomory–Hu tree on Graph $G' = (V, E \setminus E_f^*)$ with edge costs $c_h : F \to \mathbb{R}_+$. For $u, v \in V$, recall that $\mathcal{P}(u,v)$ denotes the unique path between u and v in H, and

$$\text{cut-cost}_{G'}(u,v) = \min_{e \in \mathcal{P}(u,v)} c_h(e).$$

As defined in (2.7), for any $t \in \{t_1, \ldots, t_k\}$, let

$$e_t \in \text{argmin}_{e \in \mathcal{P}(t,r)} c_h(e), \ U_t \leftarrow \text{ Component containing } t \text{ in } H \setminus \{e_t\}.$$

For any $i = 1, \ldots, k$, let

$$\begin{aligned} E_f^i \ & \leftarrow \ \delta_G(U_{t_i}) \cap E_f^*, \\ E_s^i \ & \leftarrow \ \delta_G(U_{t_i}) \setminus E_f^*. \end{aligned} \tag{3.2}$$

Note that $c_f(E_s^i) = c_f(E_s^{i*})$ for all $i = 1, \ldots, k$. Let

$$S = \{t_i \mid c_f(E_f^i) > c_f(E_s^i), i = 1, \ldots, k\}, \tag{3.3}$$

and,

$$\hat{E}_f = \delta_G\left(\bigcup_{t\in S} U_t\right).$$

Note that since $U_t, t \in S$ are laminar, as in the proof of Theorem 2.4, we can construct a minimal subset of terminals $\mathcal{T} \subseteq S$ such that

$$\bigcup_{t\in\mathcal{T}} U_t = \bigcup_{t\in S} U_t, \text{ and } U_t \cap U_{t'} = \emptyset, \ \forall t, t' \in \mathcal{T}.$$

Therefore,

$$\hat{E}_f = \delta_G\left(\bigcup_{t\in\mathcal{T}} U_t\right) = \left(\bigcup_{t\in\mathcal{T}} \delta_G(U_t)\right) \setminus E\left[\bigcup_{t\in\mathcal{T}} U_t\right].$$

Let $E^*_{f,1}, E^*_{f,2}$ be as defined in (2.11). Then similar to the proof of Theorem 2.4, we have

$$\hat{E}_f \subseteq \left(\bigcup_{t\in\mathcal{T}} \delta_G(U_t) \setminus E^*_f\right) \cup E^*_{f,1} = \left(\bigcup_{i:t_i\in\mathcal{T}} E^i_s\right) \cup E^*_{f,1}, \tag{3.4}$$

where the last equality follows from the definition of E^i_s in (3.2). Now, consider the following solution for (3.1),

$$\begin{aligned} \forall i \in [k],\ y_i \ &= \begin{cases} 1 & \text{if } t_i \in S \\ 0 & \text{otherwise} \end{cases}, \\ \forall e \in E,\ x_e \ &= \begin{cases} 1 & \text{if } e \in \hat{E}_f \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{3.5}$$

It is easy to observe that the above solution is feasible for (3.1) since the solution $x_e, e \in E$, separates terminal $t_i, i = 1, \ldots, k$ from r if $y_i = 1$. Note that $\sum_{e\in E} c_f(e) \cdot x_e = c_f(\hat{E}_f)$, and

$$\begin{aligned} c_f(\hat{E}_f) &\leq \sum_{i:t_i\in\mathcal{T}} c_f(E^i_s) + c_f(E^*_{f,1}) \\ &\leq \sum_{i:t_i\in\mathcal{T}} c_f(E^i_f) + c_f(E^*_{f,1}) \\ &= \sum_{i:t_i\in\mathcal{T}} c_f(\delta_G(U_{t_i}) \cap E^*_f) + c_f(E^*_{f,1}) \\ &= \sum_{i:t_i\in\mathcal{T}} \big(c_f(\delta_G(U_{t_i}) \cap E^*_{f,1}) + c_f(\delta_G(U_{t_i}) \cap E^*_{f,2})\big) + c_f(E^*_{f,1}) \\ &\leq\ 2 \cdot \big(c_f(E^*_{f,1}) + c_f(E^*_{f,2})\big) \\ &\leq\ 2 \cdot c_f(E^*_f) \end{aligned} \tag{3.6}$$

where the first inequality follows from (3.4) and the second inequality follows as $c_f(E_s^i) \le c_f(E_f^i)$ for all $t_i \in \mathcal{T} \subseteq S$ by the definition of S. Also,

$$\begin{aligned}\sum_{i=1}^{k} \sigma_i p_i \cdot (1 - y_i) \cdot \text{cut-cost}_G(r, t_i) &= \sum_{i: t_i \notin S} \sigma_i p_i \cdot \text{cut-cost}_G(r, t_i) \\ &\le \sum_{i: t_i \notin S} \sigma_i p_i \cdot \left(c_f(E_f^i) + c_f(E_s^i)\right) \\ &\le 2 \cdot \left(\sum_{i: t_i \notin S} \sigma_i p_i \cdot c_f(E_s^i) \right) \\ &= 2 \cdot \left(\sum_{i: t_i \notin S} \sigma_i p_i \cdot c_f(E_s^{i*})) \right),\end{aligned}$$

where the first equation follows from the solution (3.5). The second inequality follows since $E_f^i \cup E_s^i$ is an r-t_i cut in G, the third inequality follows from the definition of S (3.3) and the last equality follows as $c_f(E_s^i) = c_f(E_s^{i*})$ for all $i = 1, \ldots, k$ by construction. Combing this with (3.6), we have that the cost of solution (3.5) is at most $2z^*$ which implies $z_{IP} \le 2z^*$.■

Now, we are ready to describe our algorithm for the stochastic min-cut problem (Figure 2).

Algorithm $\mathcal{A}_{\mathsf{Stoch}}$ for Stochastic Min-Cut Problem
Input: Graph $G = (V, E)$, root r, terminals $\{t_1, t_2, \ldots, t_k\}$, inflation factors $\sigma_1, \ldots, \sigma_k$, and scenario probabilities $p_1, \ldots, p_k$.

1. Solve the LP relaxation of the integer program (3.1).
2. Let $\tilde{x}_e,\ e \in E,\ \tilde{y}_i,\ i = 1, \ldots, k$ be an optimal fractional solution for the LP relaxation.
3. Let
$$U = \left\{ t_i \;\middle|\; \tilde{y}_i \ge \frac{1}{2}, i = 1, \ldots, k \right\}.$$
4. Return

First stage solution: $E_f \leftarrow$ minimum r-U cut in G,

Second stage solution for scenario i : $E_s^i \leftarrow$ minimum r-t_i cut in $G \setminus E_f$.

Fig. 2 Algorithm for Stochastic Min-Cut Problem

Theorem 3.2 *Algorithm $\mathcal{A}_{\mathsf{Stoch}}$ gives a 4-approximation for the stochastic min-cut problem.*

Proof Let

$$\tilde{z} = \sum_{e \in E} c_f(e) \cdot \tilde{x}_e + \sum_{i=1}^{k} \sigma_i p_i \cdot (1 - \tilde{y}_i) \cdot \text{cut-cost}_G(r, t_i)$$

where $\tilde{x}_e,\ e \in E, \tilde{y}_i,\ i = 1, \ldots, k$ are an optimal solution to the LP relaxation of (3.1) as defined in $\mathcal{A}_{\mathsf{Stoch}}$. Consider the following fractional solution: $\hat{x}_e = 2\tilde{x}_e$ for all $e \in E$. For all $t \in U$,

$$\sum_{e \in P} \hat{x}_e \geq 1, \ \forall P \in \mathcal{P}(r,t).$$

Therefore, the solution, $\hat{x}$ denotes a fractional r-U cut in G which implies that

$$c_f(E_f) \leq \sum_{e \in E} c_f(e)\hat{x}_e = 2 \cdot \sum_{e \in E} c_f(e)\tilde{x}_e, \tag{3.7}$$

since E_f is a minimum r-U cut in G. Also,

$$\begin{aligned} 2da \sum_{i=1}^{k} \sigma_i p_i \cdot c_f(E_s^i) &= \sum_{i:t_i \notin U} \sigma_i p_i \cdot c_f(E_s^i) \\ &\leq \sum_{i:t_i \notin U} \sigma_i p_i \cdot \text{cut-cost}_G(r,t_i) \\ &\leq \ 2 \cdot \left(\sum_{i:t_i \notin U} \sigma_i p_i \cdot (1-\tilde{y}_i) \cdot \text{cut-cost}_G(r,t_i) \right) && (3.8) \\ &\leq 2 \cdot \left(\sum_{i=1}^{k} \sigma_i p_i \cdot (1-\tilde{y}_i) \cdot \text{cut-cost}_G(r,t_i) \right) && (3.9) \end{aligned}$$

where the first equation follows as $E_s^i = \emptyset$ for all $i : t_i \in U$. The second inequality follows as E_s^i is a r-t_i min-cut in $G \setminus E_f$ and therefore, has cost at most $\text{cut-cost}_G(r,t_i)$, and (3.8) follows since $\tilde{y}_i < 1/2$ for all $i : t_i \notin U$. Therefore,

$$\begin{aligned} c_f(E_f) + \sum_{i=1}^{k} \sigma_i p_i \cdot c_f(E_s^i) &\leq 2 \cdot \left(\sum_{e \in E} c_f(e)\tilde{x}_e + \sum_{i=1}^{k} \sigma_i p_i \cdot (1-\tilde{y}_i) \cdot \text{cut-cost}_G(r,t_i) \right) \\ &\leq 2 \cdot z_{IP} \\ &\leq 4z^*, \end{aligned}$$

where the first inequality follows from (3.7) and (3.9). The second inequality follows since $\tilde{z} \leq z_{IP}$ and the last inequality follows from Lemma 3.1. ■

4 Two-stage Robust Shortest Path Problem

In this section, we present an approximation algorithm for the two-stage robust shortest path problem, Π_{RSP} defined in (1.3) where the goal is to select a set of edges E_f (first-stage solution) and for each second-stage scenario $i = 1, \ldots, k$, edges E_s^i (second-stage solution) such that $E_f \cup E_s^i$ contain a path from r to t_i, and the objective

$$c_f(E_f) + \max_{i=1}^{k} \ \sigma_i \cdot c_f(E_s^i),$$

is minimized. Note that in the robust min-cut problem, we require that $E_f \cup E_s^i$ disconnect r from t_i.

Our algorithm is based on an approach similar to the robust min-cut problem where we guess the worst-case second-stage cost and select a subset of terminals to connect to the root in an approximate first-stage solution. The second-stage solution for any scenario i is simply a shortest path from terminal t_i to the first-stage tree. However, there are two significant differences from the robust min-cut problem. In the robust min-cut problem, terminals corresponding to all *costly* scenarios are separated from r in the first-stage, and for every other scenario i, the second-stage solution can independently pay for the r-t_i min-cut in G without using the first-stage solution. For the robust shortest path problem, however, we need to be more careful in selecting the subset of terminals to be connected to root in the first-stage solution. Moreover, for the remaining scenarios, the second-stage solution can not pay for the complete path from terminal to the root independently. Therefore, we need to analyze the cost of connecting the terminal to the approximate first-stage solution which depends on our particular choice of the first-stage tree. Due to these structural differences, the algorithm and analysis for the robust shortest path problem are significantly different from the robust min-cut problem, even though they are similar in spirit.

Let us introduce some notation before describing the algorithm. Let $\mathsf{SP}(u, v)$ denote the shortest path distance between u and v in G with respect to edge costs c_f. Let $E_f^*, E_s^{i*}, i = 1, \ldots, k$ be an optimal solution for the robust shortest path problem, and let C be the worst-case second-stage cost, i.e.,

$$C = \max_{i=1}^{k} \sigma_i \cdot c_f(E_s^{i*}).$$

We can assume that we know C; there are at most $k \cdot n$ possible choices we need to explore for C since the worst-case cost is $\sigma_i \cdot \mathsf{SP}(t_i, v)$ for some $i = 1, \ldots, k$ and $v \in V$. For any $v \in V$, $d \in \mathbb{R}_+$, let $B(v, d)$ denote the open ball of radius d around v, i.e., $B(v, d) = \{u \mid \mathsf{SP}(u, v) < d\}$. Hence, for any $u, v \in V$ we have

$$B(u, d_u) \cap B(v, d_v) = \emptyset \iff \mathsf{SP}(u, v) \geq d_u + d_v.$$

We assume that the terminals are ordered in decreasing order of inflation factors, i.e.,

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_k.$$

Fix $\alpha = 1.695$ and for all $i = 1, \ldots, k$, let $d_i = \frac{\alpha \cdot C}{\sigma_i}$. We define the set of *costly scenarios* as

$$S' = \{t_i \in T \mid \mathsf{SP}(r, t_i) > 2 \cdot d_i\}. \tag{4.1}$$

Note that for any scenario such that $t_i \notin S$, we can construct a solution in the second stage independently with cost at most $2\alpha C$. However, unlike with the minimum-cut problem, we do not compute a low-cost first-stage solution that would be feasible for all costly scenarios S' (if we did that, we would not be able to bound the first stage cost). Instead we select a subset of terminals, $S \subseteq S'$ for constructing the first-stage solution as follows: consider terminals t_i in S' in non-increasing order of inflation factors, σ_i and include it in S if $B(t_i, d_i)$ does not intersect with any $B(t_j, d_j)$ for any t_j already included in S. We describe our algorithm formally in Figure 3.

We first analyze the cost of the first-stage solution E_f computed by $\mathcal{A}_{\mathsf{RSP}}$ as compared to optimal first-stage solution E_f^*. In particular, we prove the following lemma.

Lemma 4.1 *For the first-stage solution E_f computed by $\mathcal{A}_{\mathsf{RSP}}$, we have*

$$c_f(E_f) \leq \frac{\gamma\alpha}{\alpha - 1} \cdot c_f(E_f^*),$$

where γ is the best approximation factor for the Steiner tree problem.

Algorithm $\mathcal{A}_{\mathsf{RSP}}$ for Robust Shortest Path problem

Input: Graph $G = (V, E)$, root r, terminals $T = \{t_1, t_2, \ldots, t_k\}$, inflation factors $\sigma_1 \geq \ldots \geq \sigma_k$.
Let $\alpha \leftarrow 1.695$ and $C \leftarrow$ worst-case second-stage cost of an optimal solution.

1. Let $d_i = \frac{\alpha \cdot C}{\sigma_i}$ for all $i \in [k]$, and S' as defined in (4.1).
2. Initialize $S \leftarrow \emptyset$
3. For all $t_i \in S'$ (considered in non-increasing order of σ_i)
   If $B(t_i, d_i) \cap B(t_j, d_j) = \emptyset$ for all $t_j \in S$
   $S \leftarrow S \cup \{t_i\}$.
4. $S \leftarrow S \cup \{r\}$.
5. **First stage solution:** $E_f \leftarrow$ Approximate Steiner tree on S.
6. **Second-stage solution:** $E_s^i \leftarrow$ Shortest path from t_i to E_f

Fig. 3 Algorithm for Robust Shortest Path Problem

Proof For any $t_i, t_j \in S$, $B(t_i, d_i) \cap B(t_j, d_j) = \emptyset$ which implies

$$\mathsf{SP}(t_i, t_j) \geq d_i + d_j.$$

Consider any Eulerian tour $\mathcal{T}_S$ that contains all the terminals in S. From the above inequality,

$$c_f(\mathcal{T}_S) \geq 2 \cdot \sum_{i: t_i \in S} d_i.$$

If we double the edges of an optimal Steiner tree with cost $\mathsf{St}(S)$, we get an Eulerian tour on S of cost $2 \cdot \mathsf{St}(S)$ which implies

$$\mathsf{St}(S) \geq \sum_{i: t_i \in S} d_i. \tag{4.2}$$

Consider the subgraph defined by edges, $\hat{E}$ where

$$\hat{E} = E_f^* \cup \left(\bigcup_{i: t_i \in S} E_s^{i*} \right).$$

Edges $\hat{E}$ connect all terminals in S to r and therefore, contains a Steiner tree on S. Therefore,

$$\mathsf{St}(S) \leq c_f(\hat{E}) \leq c_f(E_f^*) + \sum_{i: t_i \in S} c_f(E_s^{i*}). \tag{4.3}$$

Also, the worst-case second-stage cost of the optimal solution is C which implies that for all $i = 1, \ldots, k$,

$$c_f(E_s^{i*}) \leq \frac{C}{\sigma_i} = \frac{1}{\alpha} \cdot \frac{\alpha C}{\sigma_i} = \frac{1}{\alpha} \cdot d_i. \tag{4.4}$$

Now,

$$\begin{aligned} c_f(E_f^*) &\geq \mathsf{St}(S) - \sum_{i: t_i \in S} c_f(E_s^{i*}) \\ &\geq \mathsf{St}(S) - \sum_{i: t_i \in S} \frac{1}{\alpha} \cdot d_i && (4.5) \\ &\geq \mathsf{St}(S) - \frac{1}{\alpha} \cdot \mathsf{St}(S) && (4.6) \\ &= \left(\frac{\alpha - 1}{\alpha} \right) \cdot \mathsf{St}(S), && (4.7) \end{aligned}$$

where the first inequality follows from (4.3). Inequality (4.5) follows from (4.4) and inequality (4.6) follows from (4.2). Algorithm $\mathcal{A}_{\mathsf{RSP}}$ returns a γ-approximate Steiner tree on S as the first-stage solution, E_f. Therefore,

$$c_f(E_f) \leq \gamma \cdot \mathsf{St}(S) \leq \frac{\gamma\alpha}{\alpha - 1} \cdot c_f(E_f^*),$$

where the last inequality follows from (4.7). □

In the next lemma, we analyze the worst-case cost of the second-stage solution.

Lemma 4.2 *For the second-stage solution $E_s^i, i = 1, \ldots, k$ computed by $\mathcal{A}_{\mathsf{RSP}}$ we have*

$$\max_{i=1}^{k} \sigma_i \cdot c_f(E_s^i) \leq 2\alpha \cdot C.$$

Proof For any $i = 1, \ldots, k$, if terminal t_i does not belong to S' (defined in (4.1)),

$$\mathsf{SP}(r, t_i) \leq 2 \cdot d_i = \frac{2\alpha C}{\sigma_i} \Rightarrow \sigma_i \cdot c_f(E_s^i) \leq 2\alpha \cdot C.$$

For any terminal $t_i \in S'$, it either belongs to $S \subseteq S'$ or $S' \setminus S$. If $t_i \in S$, the first-stage solution E_f connects t_i to r and $E_s^i = \emptyset$. If $t_i \in S' \setminus S$, then there exists $t_j \in S$ such that $\sigma_j \geq \sigma_i$ and

$$B(t_i, d_i) \cap B(t_j, d_j) \neq \emptyset.$$

Therefore,

$$\mathsf{SP}(t_i, t_j) \leq d_i + d_j = \frac{\alpha C}{\sigma_i} + \frac{\alpha C}{\sigma_j} \leq \frac{2\alpha C}{\sigma_i},$$

where the last inequality follows as $\sigma_j \geq \sigma_i$. Therefore,

$$\sigma_i \cdot c_f(E_s^i) \leq \sigma_i \cdot \mathsf{SP}(t_i, t_j) \leq 2\alpha C,$$

where the first inequality follows from the fact the E_s^i is the shortest path from t_i to the first-stage Steiner tree E_f. Since E_f contains $t_j \in S$, the shortest path distance from t_i to E_f is at most the shortest path distance between t_i and t_j. □

Theorem 4.3 *Algorithm $\mathcal{A}_{\mathsf{RSP}}$ is a $(\gamma+2)$-approximation for the robust shortest path problem where γ is the best approximation factor for the Steiner tree problem.*

Proof Let $E_f, E_s^i, i = 1, \ldots, k$ denote the solution computed by $\mathcal{A}_{\mathsf{RSP}}$ and $E_f^*, E_s^{i*}, i = 1, \ldots, k$ be an optimal solution where the worst-case second-stage cost is C. Consider $\alpha = \frac{\gamma}{2} + 1$. From Lemma 4.1, we have that

$$c_f(E_f) \leq \frac{\gamma\alpha}{\alpha - 1} \cdot c_f(E_f^*) = (\gamma + 2) \cdot c_f(E_f^*).$$

From Lemma 4.2,

$$\max_{i=1}^{k} \ \sigma_i \cdot c_f(E_s^i) \leq 2\alpha \cdot C = (\gamma + 2) \cdot C.$$

Therefore, the objective value of the approximate solution is

$$\begin{aligned} c_f(E_f) + \max_{i=1}^{k} \ \sigma_i \cdot c_f(E_s^i) &\leq (\gamma + 2) \cdot \left(c_f(E_f^*) + C\right) \\ &= (\gamma + 2) \cdot \left(c_f(E_f^*) + \max_{i=1}^{k} \ \sigma_i \cdot c_f(E_s^{i*})\right). \end{aligned}$$

□

For general graphs, the best $\gamma = 1.39$ is due to Byrka et al. [4]. Thus, we have the following corollary.

Corollary 4.4 *Algorithm $\mathcal{A}_{\mathsf{RSP}}$ gives a 3.39-approximation for the robust shortest path problem.*

5 Two-stage Stochastic Shortest Path Problem

In this section, we consider the stochastic version of the shortest path problem, Π_{SSP} defined in (1.4), where the goal is to find a two-stage solution, E_f, E_s^i for all $i = 1, \ldots, k$ such that $E_f \cup E_s^i$ connects t_i to the root vertex r and the total expected cost,

$$c_f(E_f) + \sum_{i=1}^{k} p_i\sigma_i \cdot c_f(E_s^i),$$

is minimized.

In the robust version of the problem, the objective value depends only on the worst-case cost over all scenarios in the second-stage. Therefore, the algorithm is able to select a subset of terminals to connect to the root in the first-stage based on a guess for the worst-case second-stage cost. Since it can take only $k \cdot n$ possible values, the algorithm needs to try only a small number of values. However, in the stochastic version, the objective value depends on the total expected second-stage costs over all the scenarios. Therefore, the knowledge of the worst-case second-stage cost or even the total expected cost does not allow us to select a subset of scenarios to connect in the first-stage. However, we show that we can exploit an optimal fractional solution of the natural LP formulation for the problem to select a subset of scenarios whose terminals can be connected to the root in the first-stage and the expected cost in the second-stage is small.

Let us first formulate the stochastic shortest problem as a 0-1 program. For each edge $e \in E$, let $x_e \in \{0,1\}$ denote whether e is selected in the first-stage or not. Also, for all $i = 1, \ldots, k$, let $y_e^i \in \{0,1\}$ denote whether e is selected in the second-stage or not for the scenario i. For any $U \subsetneq V$, let $\delta(U)$ denote the set of edges leaving U, i.e., exactly one end point of the edge belongs to U. Now, we can formulate the stochastic shortest path problem as follows.

$$\begin{aligned} \mathsf{OPT} \;=\; \min \; & \sum_{e\in E} c_f(e)\cdot x_e + \sum_{i=1}^{k} p_i\sigma_i \cdot \sum_{e\in E} c_f(e)\cdot y_e^i \\ & \sum_{e\in\delta(U)} (x_e + y_e^i) \geq 1,\ \forall i \in [k],\ \forall U \subseteq V : t_i \in U, r \notin U \\ & x_e \in \{0,1\},\ \forall e \in E \\ & y_e^i \in \{0,1\},\ \forall e \in E, \forall i = 1 \in [k]. \end{aligned} \tag{5.1}$$

We consider the LP relaxation of (5.1) where we relax the integrality constraints on x_e, y_e^i for all $e \in E, i \in [k]$.

$$
\begin{aligned}
z_{\mathsf{SSP}}^{\mathsf{LP}} \;=\; \min \; & \sum_{e\in E} c_f(e)\cdot x_e + \sum_{i=1}^{k} p_i\sigma_i \cdot \sum_{e\in E} c_f(e)\cdot y_e^i \\
& \sum_{e\in\delta(U)} (x_e + y_e^i) \geq 1, \;\; \forall i\in[k],\; \forall U\subseteq V : t_i\in U, r\notin U \qquad (5.2)\\
& x_e \geq 0 \;\forall e\in E \\
& y_e^i \geq 0, \;\forall e\in E, \forall i = 1\in[k].
\end{aligned}
$$

Note that the above LP has exponentially many constraints but the separation problem and therefore, the optimization problem can be solved in polynomial time. The separation problem is the following: given a solution x_e, y_e^i for all $e \in E$, $i \in [k]$, we need to decide whether the cut-constraints are satisfied for all $i = 1, \ldots, k$. For each $i = 1, \ldots, k$, consider edge costs $c^i : E \to \mathbb{R}_+$ where $c^i(e) = x_e + y_e^i$ for all $e \in E$. Compute

$$\theta_i := r\text{-}t_i \text{ min-cut in graph } G \text{ with respect to edge costs } c^i.$$

Then the solution is feasible if and only if $\theta_i \geq 1$ for all $i \in [k]$. If for some scenario i, $\theta_i < 1$, the corresponding cut gives a violated inequality.

Let x_e^*, y_e^{i*} for $e \in E, i \in [k]$ be an optimal LP solution for (5.2). Let $\alpha = 3.39$, and

$$d_i = \alpha \cdot \sum_{e\in E} c_f(e)\cdot y_e^{i*}, \;\forall i = 1,\ldots,k. \qquad (5.3)$$

We assume that the terminals are ordered in non–decreasing order of d_i, i.e.,

$$d_1 \leq d_2 \leq \ldots \leq d_k.$$

As earlier, for any $u, v \in V$, $\mathsf{SP}(u, v)$ denotes the shortest path distance between u and v in G, $B(v, d)$ denotes a ball of radius $d \in \mathbb{R}_+$ around v, and

$$B(u, d_u) \cap B(v, d_v) = \emptyset \iff \mathsf{SP}(u, v) \geq d_u + d_v.$$

We incrementally construct a subset $S \subseteq T$ of terminals that are connected to r in a first-stage solution. We consider each terminal t_i in non–decreasing order of d_i, and include it in S if $B(t_i, d_i)$ does not intersect with $B(t_j, d_j)$ for any t_j that already belongs to S. We describe our algorithm formally in Figure 4.

We first analyze the cost of the first-stage solution E_f computed by $\mathcal{A}_{\mathsf{SSP}}$. In particular, we prove the following lemma.

Lemma 5.1 *Let x_e^*, y_e^{i*} for all $e \in E, i \in [k]$ be an optimal LP solution for* (5.2)*, and E_f be the first-stage solution computed by $\mathcal{A}_{\mathsf{SSP}}$. Then*

$$c_f(E_f) \leq \left(\frac{2\gamma\alpha}{\alpha - 2}\right) \cdot \sum_{e\in E} c_f(e)\cdot x_e^*,$$

where γ is the best approximation factor for the Steiner tree problem.

Algorithm, $\mathcal{A}_{\mathsf{SSP}}$ for Stochastic Shortest Path problem

Input: $G = (V, E)$, root r, $T = \{t_1, t_2, \ldots, t_k\}$, $\sigma_1, \ldots, \sigma_k, p_1, \ldots, p_k$.
Let $\alpha \leftarrow 3.39$.

1. Solve LP (5.2) and let x_e^*, y_e^{i*}, $e \in E, i \in [k]$ be an optimal solution.
2. For all $i = 1, \ldots, k$, let $d_i = \alpha \cdot \sum_{e \in E} c_f(e) \cdot y_e^{i*}$.
3. Reorder the terminals in non–decreasing order of d_i.
4. Initialize $S \leftarrow \emptyset$
5. For all $i = 1$ to k
   If $B(t_i, d_i) \cap B(t_j, d_j) = \emptyset$ for all $t_j \in S$
      $S \leftarrow S \cup \{t_i\}$.
6. $S \leftarrow S \cup \{r\}$.
7. **First stage solution:** $E_f \leftarrow$ Approximate Steiner tree on S.
8. **Second-stage solution:** $E_s^i \leftarrow$ Shortest path from t_i to E_f

Fig. 4 Algorithm for Stochastic Shortest Path Problem

Proof Using an argument similar to the proof of Lemma 4.1, we can show that

$$\mathsf{St}(S) \geq \sum_{i:t_i \in S} d_i = \sum_{i:t_i \in S} \alpha \cdot \sum_{e \in E} c_f(e) \cdot y_e^{i*}, \tag{5.4}$$

where $\mathsf{St}(S)$ denotes the cost of optimal Steiner tree on S. By construction $B(t_i, d_i) \cap B(t_j, d_j) = \emptyset$ for any $t_i, t_j \in S$ which implies

$$\mathsf{SP}(t_i, t_j) \geq d_i + d_j.$$

If we double the edges of an optimal Steiner tree with cost $\mathsf{St}(S)$, we get an Eulerian tour on S of cost $2 \cdot \mathsf{St}(S)$ which implies (5.4).

Consider the following fractional solution,

$$z_e = x_e^* + \sum_{i:t_i \in S} y_e^{i*}, \ \forall e \in E.$$

We claim that $z_e, e \in E$ is a feasible solution for the Steiner tree LP on terminals S. Consider any $U \subseteq V$ that separates terminals in S, i.e., $S \cap U \neq \emptyset$ and $S \setminus U \neq \emptyset$. It is easy to observe that $z_e, e \in E$ satisfies the cut inequality corresponding to U. Since U separates terminals in S, U must separate some terminal, say $t_j \in S$ from r. Therefore,

$$\begin{aligned}\sum_{e \in \delta(U)} z_e &= \sum_{e \in \delta(U)} \left(x_e^* + \sum_{i:t_i \in S} y_e^{i*} \right) \\ &\geq \sum_{e \in \delta(U)} x_e^* + y_e^{j*} \\ &\geq 1,\end{aligned}$$

where the last inequality follows since $x_e^*, y_e^{i*}, e \in E, i \in [k]$ is feasible for (5.2). From Agarwal et al. [1] and Goemans and Williamson [8], we know that

$$\mathsf{St}(S) \leq 2 \cdot \sum_{e \in E} c_f(e) \cdot z_e = 2 \cdot \left(\sum_{e \in E} c_f(e) \cdot x_e^* + \sum_{i:t_i \in S} \sum_{e \in E} c_f(e) \cdot y_e^{i*} \right). \tag{5.5}$$

Therefore,

$$\begin{aligned}\sum_{e\in E} c_f(e)\cdot x_e^* &\geq \frac{1}{2}\cdot \mathsf{St}(S) - \sum_{i:t_i\in S}\sum_{e\in E} c_f(e)\cdot y_e^{i*} \\ &\geq \frac{1}{2}\cdot \mathsf{St}(S) - \frac{1}{\alpha}\cdot \mathsf{St}(S) \qquad (5.6)\\ &= \left(\frac{\alpha-2}{2\alpha}\right)\cdot \mathsf{St}(S),\end{aligned}$$

where (5.6) follows from (5.4). Algorithm $\mathcal{A}_{\mathsf{SSP}}$ returns a γ-approximate Steiner tree on S as the first-stage solution, E_f. Therefore,

$$c_f(E_f) \leq \gamma\cdot \mathsf{St}(S) \leq \left(\frac{2\gamma\alpha}{\alpha-2}\right)\cdot \sum_{e\in E} c_f(e)\cdot x_e^*,$$

where the last inequality follows from (5.6). □

In the next lemma, we analyze the expected second-stage cost. In particular, we prove the following.

Lemma 5.2 *Let E_s^i denote the second-stage solution for scenario $i=1,\ldots,k$ computed by $\mathcal{A}_{\mathsf{SSP}}$. Then*

$$\sum_{i=1}^{k} p_i\sigma_i\cdot c_f(E_s^i) \leq 2\alpha\cdot\left(\sum_{i=1}^{k} p_i\sigma_i\cdot\sum_{e\in E} c_f(e) y_e^{i*}\right).$$

Proof For any $i=1,\ldots,k$, if terminal $t_i\in S$, t_i is connected to r in the first-stage and $E_s^i=\emptyset$. If $t_i\notin S$, then there exist some $t_j\in S$ with $d_j\leq d_i$ such that $B(t_i,d_i)\cap B(t_j,d_j)\neq\emptyset$ which implies

$$\mathsf{SP}(t_i,t_j)\leq d_i+d_j\leq 2d_i.$$

The second stage solution, E_s^i for scenario i is a shortest path from t_i to the first-stage Steiner tree E_f that contains t_j. Therefore,

$$c_f(E_s^i)\leq \mathsf{SP}(t_i,t_j)\leq 2d_i, \qquad (5.7)$$

Now,

$$\begin{aligned}\sum_{i=1}^{k} p_i\sigma_i\cdot c_f(E_s^i) &= \sum_{i:t_i\notin S} p_i\sigma_i\cdot c_f(E_s^i)\\ &\leq \sum_{i:t_i\notin S} p_i\sigma_i\cdot (2d_i) \qquad (5.8)\\ &= 2\alpha\cdot\left(\sum_{i:t_i\notin S} p_i\sigma_i\cdot\sum_{e\in E} c_f(e) y_e^{i*}\right)\\ &\leq 2\alpha\cdot\left(\sum_{i=1}^{k} p_i\sigma_i\cdot\sum_{e\in E} c_f(e) y_e^{i*}\right),\end{aligned}$$

where (5.8) follows from (5.7). □

Theorem 5.3 *Algorithm $\mathcal{A}_{\mathsf{SSP}}$ is a $2(\gamma+2)$-approximation for the stochastic shortest path problem where γ is the best approximation factor for the Steiner tree problem.*

Proof Let $E_f, E_s^i, i = 1, \ldots, k$ be the solution computed by $\mathcal{A}_{\mathsf{SSP}}$ and x_e^*, y_e^{i*} for all $e \in E, i \in [k]$ be an optimal solution LP solution for (5.2). Consider $\alpha = \gamma+2$. From Lemma 5.1, we have

$$c_f(E_f) \leq \left(\frac{2\gamma\alpha}{\alpha-1}\right) \cdot \sum_{e\in E} c_f(e) \cdot x_e^*,$$

and from Lemma 5.2,

$$\sum_{i=1}^{k} p_i\sigma_i \cdot c_f(E_s^i) \leq 2\alpha \cdot \left(\sum_{i=1}^{k} p_i\sigma_i \cdot \sum_{e\in E} c_f(e) y_e^{i*}\right).$$

Therefore, the objective value of the approximate solution is

$$\begin{aligned} c_f(E_f) + \sum_{i=1}^{k} p_i\sigma_i \cdot c_f(E_s^i) &\leq \frac{2\gamma\alpha}{\alpha-1} \cdot \left(\sum_{e\in E} c_f(e) x_e^*\right) + 2\alpha \cdot \left(\sum_{i=1}^{k} p_i\sigma_i \cdot \sum_{e\in E} c_f(e) y_e^{i*}\right) \\ &\leq 2(\gamma+2) \cdot \left(\sum_{e\in E} c_f(e) \cdot x_e^* + \sum_{i=1}^{k} p_i\sigma_i \cdot \sum_{e\in E} c_f(e) y_e^{i*}\right) \\ &= 2(\gamma+2) \cdot z_{\mathsf{SSP}}^{\mathsf{LP}} \\ &\leq 2(\gamma+2) \cdot \mathsf{OPT}. \end{aligned}$$

where the last inequality follows as $z_{\mathsf{SSP}}^{\mathsf{LP}} \leq \mathsf{OPT}$. □

For general graphs, the best $\gamma = 1.39$ is due to Byrka et al. [4]. Thus, we have the following corollary.

Corollary 5.4 *Algorithm $\mathcal{A}_{\mathsf{SSP}}$ gives a 6.78-approximation for the stochastic shortest path problem.*

6 Conclusions

In this paper, we present improved approximations for the robust and stochastic min-cut and shortest path problems. In particular, we give a 2-approximation and 4-approximation for the robust and stochastic min-cut problems respectively, and 3.39-approximation and 6.78-approximation for the robust and stochastic shortest path problems respectively. Our algorithms for the robust versions of the problem are based on a guess and prune strategy, namely, we guess the worst-case second-stage cost of an optimal solution and use it to select (or prune) a subset of costly scenarios that are completely satisfied in the first-stage. The algorithms for both the robust min-cut and robust shortest path problem utilize a crucial property of the robust objective: the objective value depends only on the worst-case cost in the second-stage, and therefore, each scenario can pay up to the worst-case cost in the second-stage. This property allow us to use the guess and prune strategy for the robust problems.

For the stochastic versions, the objective is to minimize the total expected cost and not the worst-case cost. Therefore, the guess and prune algorithms do not directly apply to the stochastic versions directly. Interestingly, we use LP formulations to adapt the guess and prune algorithm of the robust versions for both stochastic min-cut and shortest path problems. However, the approximation factors get worse by a factor of two as compared to the robust versions. It would be interesting to explore if the techniques in this paper can be extended to robust and stochastic versions of Steiner tree and multi-cut problems where each scenario contains a set of terminals instead of just a single terminal in shortest path and min-cut problems.

References

1. A. Agrawal, P. Klein, and R. Ravi. When trees collide: An approximation algorithm for the generalized Steiner problem on networks. *SIAM Journal on Computing*, 24(3):440–456, 1995.
2. A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust optimization*. Princeton University press, 2009.
3. D. Bertsimas, D.B. Brown, and C. Caramanis. Theory and applications of Robust Optimization. *SIAM Review (to Appear)*, 2009.
4. Jaroslaw Byrka, Fabrizio Grandoni, Thomas Rothvoß, and Laura Sanità. An improved lp-based approximation for Steiner tree. In *STOC '10: Proceedings of the 42nd ACM symposium on Theory of computing*, pages 583–592, New York, NY, USA, 2010. ACM.
5. M. Charikar, C. Chekuri, and M. Pál. Sampling bounds for stochastic optimization. *Approximation, Randomization and Combinatorial Optimization. Algorithms and Techniques*, pages 610–610, 2005.
6. K. Dhamdhere, V. Goyal, R. Ravi, and M. Singh. How to pay, come what may: Approximation algorithms for demand-robust covering problems. In *FOCS*, pages 367–378, 2005.
7. U. Feige, K. Jain, M. Mahdian, and V. Mirrokni. Robust combinatorial optimization with exponential scenarios. *Integer Programming and Combinatorial Optimization*, pages 439–453, 2007.
8. M.X. Goemans and D.P. Williamson. A general approximation technique for constrained forest problems. *SIAM Journal on Computing*, 24(2):296–317, 1995.
9. A. V. Goldberg and R. E. Tarjan. A new approach to the maximum-flow problem. *J. ACM*, 35(4):921–940, 1988.
10. D. Golovin, V. Goyal, and R. Ravi. Pay today for a rainy day: Improved approximations for demand-robust min-cut and shortest path problems. In *STACS*, pages 206–217, 2006.
11. Gomory, R. E. and Hu, T. C. Multi-terminal network flows. *Journal of the Society for Industrial and Applied Mathematics*, 9(4):551–570, dec 1961.
12. A. Gupta, V. Nagarajan, and R. Ravi. Thresholded covering algorithms for robust and max-min optimization. *Automata, Languages and Programming*, pages 262–274, 2010.
13. A. Gupta, M. Pál, R. Ravi, and A. Sinha. Boosted sampling: approximation algorithms for stochastic optimization. In *STOC*, pages 417–426, 2004.
14. 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 684–693, 2004.
15. G. Infanger. *Planning under uncertainty: solving large-scale stochastic linear programs*. Boyd & Fraser Pub Co, 1994.
16. P. Kall and S.W. Wallace. *Stochastic programming*. Wiley New York, 1994.
17. R. Khandekar, G. Kortsarz, V. Mirrokni, and M. Salavatipour. Two-stage robust network design with exponential scenarios. *Algorithms-ESA 2008*, pages 589–600, 2008.
18. A. Prékopa. *Stochastic programming*. Kluwer Academic Publishers, Dordrecht, Boston, 1995.
19. R. Ravi and A. Sinha. Hedging uncertainty: Approximation algorithms for stochastic optimization problems. In *IPCO*, pages 101–115, 2004.
20. A. Shapiro. Stochastic programming approach to optimization under uncertainty. *Mathematical Programming, Series B*, 112(1):183–220, 2008.
21. A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on stochastic programming: modeling and theory*. Society for Industrial and Applied Mathematics, 2009.
22. D.B. Shmoys and C. Swamy. Stochastic optimization is (almost) as easy as deterministic optimization. In *Foundations of Computer Science, 2004. Proceedings. 45th Annual IEEE Symposium on*, pages 228–237. IEEE, 2004.