MANAGEMENT SCIENCE
Vol. 56, No. 10, October 2010, pp. 1794–1814
ISSN 0025-1909 | EISSN 1526-5501 | 10 | 5610 | 1794

DOI 10.1287/mnsc.1100.1208

Expectation and Chance-Constrained Models and Algorithms for Insuring Critical Paths

Siqian Shen, J. Cole Smith
Department of Industrial and Systems Engineering, University of Florida, Gainesville, Florida 32611
{sshen@ufl.edu, cole@ise.ufl.edu}

Shabbir Ahmed
H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332, sahmed@isye.gatech.edu

In this paper, we consider a class of two-stage stochastic optimization problems arising in the protection of vital arcs in a critical path network. A project is completed after a series of dependent tasks are all finished. We analyze a problem in which task finishing times are uncertain but can be insured a priori to mitigate potential delays. A decision maker must trade off costs incurred in insuring arcs with expected penalties associated with late project completion times, where lateness penalties are assumed to be lower semicontinuous nondecreasing functions of completion time. We provide decomposition strategies to solve this problem with respect to either convex or nonconvex penalty functions. In particular, for the nonconvex penalty case, we employ the reformulation-linearization technique to make the problem amenable to solution via Benders decomposition. We also consider a chance-constrained version of this problem, in which the probability of completing a project on time is sufficiently large. We demonstrate the computational efficacy of our approach by testing a set of size-and-complexity diversified problems, using the sample average approximation method to guide our scenario generation.

Key words: project management; integer programming; reformulation-linearization technique; chance-constrained programming; sample average approximation
History: Received May 10, 2009; accepted March 8, 2010, by Dimitris Bertsimas, optimization. Published online in *Articles in Advance* August 3, 2010.

1. Introduction

In this paper, we analyze a class of critical path management (CPM) problems associated with the scheduling of complex projects that consist of several dependent activities. These problems are often modeled by a directed network whose arcs represent dependent tasks and whose arc weights consist of associated task durations. Graph vertices represent milestones, and they enforce precedence constraints by requiring that all tasks associated with arcs entering a node must be completed before any tasks associated with arcs exiting the node may begin. Also, there exists a node representing the start of the project and a node representing its termination. If the duration for each activity is not known with certainty, the program evaluation and review technique (PERT) can be used to estimate the probability that a project will be completed by a given deadline. (See, e.g., Kelley 1961, 1963; Moehring 1984 for basic CPM and PERT literature.) Given the CPM network associated with a project, the total project completion time is given by the length of the network's longest path, referred to as its critical path. Critical path lengths are carefully monitored in most business applications because financial penalties often accrue as a monotonic function of project completion time.

Some recent models and algorithms for resource-constrained project scheduling problems include works by Brucker et al. (1999), Chtourou and Haouari (2008), Ozadamar and Ulusoy (1995), and Patterson (1984). In particular, Elamaghraby et al. (2000), Hagstrom (1990), and Iida (2000) compute lower and upper bounds for the distribution function of PERT project durations in order to schedule tasks. From a multistage dynamic perspective, Bowman and Muckstadt (1993), Hindelang and Muth (1979), and Kulkarni and Adlakha (1986) employ dynamic programming and finite-stage, continuous-time Markov chains to solve these problems.

Several methodologies are used to treat mass uncertain information, including heuristic-based and Monte Carlo simulation-based techniques (Burt and Garman 1971, Bowman 1995, Mitchell and Klastorin 2007). Critical chain project management (CCPM) (Goldratt 1997) is a method based on theory of constraints that emphasizes keeping resources flexible at start times and quickly switching resources between tasks as necessary. Herroelen and Leus (2001) highlight the merits and disadvantages of CCPM based on literature and experimental studies on commercial CCPM software. By contrast, in this paper, we consider situations in which resources cannot be dynamically switched among tasks during the execution of the project. These situations occur, for instance, when resources must be allocated well in advance (e.g., in capital budgeting plans) or when communication/transportation logistics make resource transfer impossible.

Herroelen and Leus (2005) review fundamental approaches for project management under uncertainty, including reactive scheduling, stochastic project scheduling, fuzzy project scheduling, and robust (proactive) scheduling. The *time/cost trade-off* (or *activity crashing*) problems are related to the study in this paper. Scholl (2001) and Gutjahr et al. (2000) formulate an expectation-based version of the stochastic linear time/cost trade-off problem as a scenario-based stochastic program. These problems have also recently been analyzed using chance-constrained formulations (Laslo 2003, Golenko-Ginzburg and Gonik 1998, Golenko-Ginzburg et al. 2000). We refer to Demeulemeester and Herroelen (2002) for a comprehensive discussion of contemporary stochastic project scheduling problems.

We consider the cases in which project managers can invest resources to either shorten task durations or prevent spikes in task durations because of uncertainty. For instance, task durations may be substantially delayed because of labor availability in construction or shipping delays in logistics applications. In these settings, insuring tasks against delays may be accomplished by prehiring additional labor to keep in reserve or by paying additional money to guarantee timely delivery of goods. A decision maker would seek an optimal portfolio of resource investment, trading off costs of insuring arcs with expected penalties associated with project deadline violation. We refer to this problem as the stochastic task insurance problem (STIP).

Note that the STIP is also related to interdiction problems, which is another class of two-stage problems that often take place over networks. A typical network interdiction problem involves a network operator that wishes to minimize (without loss of generality) some objective over the network, such as a shortest path or minimum cost flow. The interdiction problem is set up as a Stackelberg game wherein an interdicting agent acts first to modify the characteristics of certain arcs (e.g., reducing or eliminating capacity) in order to maximize the operator's minimum cost. Some stochastic interdiction problems of note include those by Cormican et al. (1998) and Janjarassuk and Linderoth (2008).

In this paper, we formulate the STIP as a two-stage stochastic programming model amenable to Benders decomposition (Benders 1962). (See Schultz 2003 for a review of stochastic integer programming models and algorithms, and see Chen et al. 2008 for multistage stochastic optimization models.) Our primary contributions in this paper are as follows. First, we propose a Benders decomposition framework for the solution of STIP in which lateness is penalized by a nondecreasing lower semicontinuous penalty function of project completion time. These functions are of significant practical importance, because they allow a decision maker to capture discontinuities and/or smooth nonconvex portions of the penalty function. Discontinuities may arise because of fixed-charge fees due to lateness, and smooth nonconvexities may arise when penalty functions are concave functions that asymptotically approach a maximum value (e.g., because of project cancellation). Second, we demonstrate how to quickly recover coefficients for Benders cuts from the solution of a critical path problem, rather than requiring the direct solution of a more complex reformulation. Third, we cast the STIP in the context of a chance-constrained optimization problem and demonstrate how our algorithms can be used to solve such instances.

We then conduct a computational study in §4 that both illustrates the efficiency of our procedures and demonstrates the inherent difficulty of solving the STIP. In particular, we examine two intuitive methods that managers may be tempted to use to determine which tasks should be insured. In §4.1.2, we consider the use of a simple rule in which tasks are insured in order of a nondecreasing ratio of task insurance cost to average time saved because of insurance. In §4.1.3, we solve a series of deterministic critical path insurance problems, in which the scenario outcomes are known a priori, to obtain an empirical estimate of how likely a task is to be insured in each scenario. (We refer to this as the *persistency* of a task; see Bertsimas et al. 2006.) Hence, another intuitive statement may suppose that tasks having high persistencies are more likely to be insured at optimality in the STIP. However, we demonstrate that neither of these approaches are capable of reliably picking optimal tasks to insure in the STIP. The vital implication is that the STIP is too difficult to be solved by examining cost-to-benefit ratios as supposed in §4.1.2, and is even too difficult to be solved by insuring those tasks that appear most often in the solution to a series of deterministic task insurance problems in which uncertain information is revealed before the insurance decisions take place.

Finally, note that the actual completion time of the project will be a function of the task insurance decisions and the outcome of the random task durations, and also of the penalty function that a manager may place on late completion times. We consider in §4.1.4 the case in which a decision maker enforces a continuous two-segment piecewise-linear penalty function on the late completion times. We investigate the effect of concavity and convexity of this penalty function on the completion time distribution. We observe that with convex penalties the average critical path length tends to be shorter than with concave penalties, explained by the fact that in the former case, severe penalty is imposed on very late completion times. This observation is of particular interest to managers that are risk averse and wish to mitigate worst-case scenarios.

The remainder of this paper is organized as follows. In §2, we introduce our problem and provide a subgradient-based cutting-plane algorithm with respect to the convex penalty case. In §3, we first examine the case of piecewise-linear lower semicontinuous penalty functions, for which we employ the reformulation-linearization technique (RLT) of Sherali and Adams (1990, 1994) to remodel the subproblem so that it is amenable to solution via Benders decomposition. We then extend our algorithm to handle general nondecreasing lower-semicontinuous penalty functions, and also solve a variation of the STIP that we cast as a chance-constrained formulation. In §4, we employ the sample average approximation (SAA) method (see, e.g., Shapiro and Homem-de-Mello 2000) to solve instances having stochastic task durations. Finally, we state our conclusions in §5.

2. Problem Statement and Convex Penalty Case

Let $G(\mathcal{N}, \mathcal{A})$ denote a directed graph representing the tasks to be completed in a complex project, with node set $\mathcal{N} = \{0, \ldots, n\}$, and arc set $\mathcal{A} \subset \mathcal{N} \times \mathcal{N}$, where $\mathcal{A}$ is topologically ordered such that $(i, j) \in \mathcal{A}$ only if $i < j$. Node 0 serves as the project starting point and node n as its completion point. We define $FS(i) = \{j\colon (i, j) \in \mathcal{A}\}$ as the set of nodes adjacent from node i, and $RS(i) = \{j\colon (j, i) \in \mathcal{A}\}$ as the set of nodes adjacent to node i, $\forall\, i \in \mathcal{N}$.

For each arc $(i, j) \in \mathcal{A}$, we represent the cost of insuring (i, j) by c_{ij}, and define binary decision variable x_{ij}, where $x_{ij} = 1$ if we insure arc (i, j) and $x_{ij} = 0$ otherwise. The set of possible finite scenarios is given by Ω, where for each scenario $s \in \Omega$, each arc $(i, j) \in \mathcal{A}$ is associated with an uninsured task duration d_{ij}^s and insured task duration g_{ij}^s, where $g_{ij}^s \in [0, d_{ij}^s]$. We use binary variables y_{ij}^s to denote whether arc (i, j) belongs to a critical path in scenario $s \in \Omega$, where $y_{ij}^s = 1$ if arc (i, j) is part of one identified critical path.

For our initial model, suppose that we have nondecreasing convex functions $\Theta^s\colon \mathbb{R} \mapsto \mathbb{R}$ that penalize the critical path length in each scenario. We define e^s as the probability of realizing a scenario $s \in \Omega$, and present the optimization problem as

$$\textbf{CP:}\quad \min\left\{\sum_{(i,j)\in\mathcal{A}} c_{ij}x_{ij} + \sum_{s\in\Omega} e^s \Theta^s(\psi^s(x))\right\}$$
$$\text{subject to: } x_{ij} \in \{0, 1\} \quad \forall\, (i, j) \in \mathcal{A},$$

where $\Theta^s(\psi^s(x))$ is the penalty function in scenario $s \in \Omega$, and $\psi^s(x)$ is the critical path length with respect to x, given by

$$\textbf{CPM}^s(x)\textbf{:}\quad \max \sum_{(i,j)\in\mathcal{A}} (d_{ij}^s - (d_{ij}^s - g_{ij}^s)x_{ij})y_{ij}^s \tag{1}$$
$$\text{subject to: } \sum_{j\in FS(0)} y_{0j}^s = 1 \tag{2}$$
$$\sum_{j\in FS(i)} y_{ij}^s - \sum_{k\in RS(i)} y_{ki}^s = 0 \quad \forall\, i \in \mathcal{N} - \{0, n\} \tag{3}$$
$$0 \le y_{ij}^s \le 1 \quad \forall\, (i, j) \in \mathcal{A}, \tag{4}$$

where (1) maximizes the sum of task durations, (2) and (3) enforce flow-balance constraints for critical path contiguity, and (4) bounds the y variables between 0 and 1.

Because $\psi^s(x)$ is convex in x and Θ^s is a nondecreasing convex function of $\psi^s(x)$, we have that $f^s(x) = \Theta^s(\psi^s(x))$ is convex in x. Letting η_s denote the optimal objective value of the subproblem based on scenario s, a standard Benders decomposition approach would create the following (relaxed) master problem:

$$\textbf{CP-MP:}\quad \min\left\{\sum_{(i,j)\in\mathcal{A}} c_{ij}x_{ij} + \sum_{s\in\Omega} e^s \eta_s\right\}$$
$$\text{subject to: } \eta_s \ge f^s(x^t) + [\partial f^s(x^t)]^T (x - x^t) \quad \forall\, t \in \mathbb{T}^s,\ s \in \Omega \tag{5}$$
$$x_{ij} \in \{0, 1\} \quad \forall\, (i, j) \in \mathcal{A},$$

where $\mathbb{T}^s$ is the collection of optimality cuts under scenario s, x^t is the candidate solution in the tth iteration, and $\partial f^s(x^t)$ is a subgradient of $f^s(x)$ at x^t.

To compute the corresponding parameters in (5), rather than requiring the direct dual solution of a subproblem reformulation, note that $f^s(x^t)$ is directly obtainable from solving a critical path problem $\text{CPM}^s(x^t)$ and setting $f^s(x^t) = \Theta^s(\psi^s(x^t))$. Let $\Theta^{s,t}$ be the slope of $\Theta^s(\cdot)$ at $\psi^s(x^t)$, and given x^t, let $y^{s,t}$ be the optimal solution of $\text{CPM}^s(x^t)$. We have that $\{-(d_{ij}^s - g_{ij}^s)y_{ij}^{s,t}\}$ is the (i, j)th element of a subgradient of $\psi^s(x)$ at x^t, and so

$$\partial f_{ij}^s(x^t) = -\Theta^{s,t}(d_{ij}^s - g_{ij}^s)y_{ij}^{s,t} \quad \forall\, (i, j) \in \mathcal{A}. \tag{6}$$

3. Decomposition Algorithm for the Nonconvex Penalty Case

We analyze a class of problems involving nonconvex penalty functions, where the cuts generated using the subgradient method in §2 are no longer valid with respect to the nonconvex subproblems. For piecewise-linear lower semicontinuous penalty functions, we employ RLT to convexify the second-stage programs, and generate Benders cuts associated with the modified subproblems in §3.1. We then extend these techniques to handle more general nonconvex functions in §3.2, and employ our algorithms to solve a chance-constrained formulation of our problem in §3.3.

3.1. Piecewise-Linear Lower Semicontinuous Penalty Function

We begin by considering STIPs in which the completion time penalty is given by a piecewise-linear lower semicontinuous function, and we develop a modified Benders decomposition method for the problem. We decompose STIP as a two-stage stochastic mixed-integer program that has binary x variables and mixed-integer recourse variables in the first and second stages, respectively, and independent subproblems for each scenario $s \in \Omega$.

We formulate the master problem as

$$\textbf{MP-PW:} \quad \min\left\{\sum_{(i,j)\in\mathcal{A}} c_{ij}x_{ij} + \sum_{s\in\Omega} e^s\eta_s\right\} \tag{7}$$

$$\text{subject to:} \quad \eta_s + m^{s,l}x \geq n^{s,l} \quad \forall\, l \in L^s,\ s \in \Omega \tag{8}$$

$$x_{ij} \in \{0,1\} \quad \forall\,(i,j) \in \mathcal{A}$$

$$\eta_s \geq \underline{\eta}^s \quad \forall\, s \in \Omega, \tag{9}$$

where L^s designates a set of Benders cuts derived for the sth subproblem, having coefficients $m^{s,l}$ and $n^{s,l}$, as we describe below, and $\underline{\eta}^s$ denotes the smallest penalty that could be incurred in scenario s for any choice of x. (Note that the penalty associated with the critical path length when all task durations are set to the g^s values is a valid lower bound for $\underline{\eta}^s$.)

3.1.1. Subproblem Formulation and Solving Algorithms. For scenario $s \in \Omega$ we define Θ^s over intervals $1, \ldots, K^s$, where interval k is defined over $(\tau_k^s, \tau_{k+1}^s]$, for $k = 1, \ldots, K^s$, and where $\tau_{K^s+1}^s$ is a maximum possible critical path length in scenario s. For convenience, we assume that τ_1^s is just smaller than $\underline{\eta}^s$ to allow all intervals to be open on the left side. For each scenario $s \in \Omega$, the piecewise-linear penalty function has slope m_k^s and intercept b_k^s over interval k, $\forall\, k = 1, \ldots, K^s$, where $m_k^s \geq 0$, $\forall\, k$, and $b_1^s \geq 0$ to ensure that each function is nonnegative and nondecreasing.

We define binary variable z_k^s such that $z_k^s = 1$ if the project completion time belongs to interval k, and $z_k^s = 0$ otherwise, for $k = 1, \ldots, K^s$ and $s \in \Omega$. Define u_i^s to represent the length of a longest path from node 0 to node i, $\forall\, i = 0, \ldots, n$, $s \in \Omega$, given the values of $\hat{x}_{ij}$ from the first-stage master problem (where $u_0^s = 0$). Also, define variables f_k^s as the objective function contribution due to interval k in scenario s, i.e., $f_k^s = (m_k^s u_n^s + b_k^s)z_k^s$. Letting Q denote an arbitrarily large number, we formulate the subproblem as

$$\textbf{SP-LS}^s(\hat{x}): \quad \min \sum_{k=1}^{K^s} f_k^s \tag{10}$$

$$\text{subject to:} \quad f_k^s \geq m_k^s u_n^s + b_k^s - Q(1 - z_k^s) \quad \forall\, k = 1, \ldots, K^s \tag{11}$$

$$u_n^s \geq \tau_k^s z_k^s \quad \forall\, k = 1, \ldots, K^s \tag{12}$$

$$-u_n^s \geq -\tau_{k+1}^s - Q(1 - z_k^s) \quad \forall\, k = 1, \ldots, K^s \tag{13}$$

$$u_j^s - u_i^s \geq d_{ij}^s - (d_{ij}^s - g_{ij}^s)\hat{x}_{ij} \quad \forall\,(i,j) \in \mathcal{A} \tag{14}$$

$$f_k^s \geq 0 \quad \forall\, k = 1, \ldots, K^s \tag{15}$$

$$\sum_{k=1}^{K^s} z_k^s = 1 \tag{16}$$

$$z_k^s \in \{0,1\} \quad \forall\, k = 1, \ldots, K^s. \tag{17}$$

We employ RLT to reformulate the STIP into a form that is amenable to solution by Benders decomposition. (See also Sherali 2001 for related development on applying RLT to piecewise-linear lower semicontinuous functions.) By noticing that $\sum_{k=1}^{K^s} z_k^s = 1$, $\forall\, s \in \Omega$, we generate the convex hull of solutions to the overall STIP (containing all scenarios) in which z_k^s is binary valued $\forall\, k, s$, based on the special structures RLT of Sherali et al. (1998), and the risk-based decomposition strategy of Sherali and Smith (2009). (See also Sherali and Fraticelli 2002 for related work.) The key is to multiply (11)–(16), as well as the bounds $x_{ij} \leq 1$ and $x_{ij} \geq 0$, by z_k^s, $\forall\, k = 1, \ldots, K^s$, and $x_{ij} \geq 0$ by $(1 - \sum_{k=1}^{K^s} z_k^s)$ (yielding an equality constraint), before decomposition. After doing so, there exists an optimal solution in which all z variables are binary valued, given binary x values, and the problem can thus be solved by Benders decomposition.

Observe that, by definition, $f_k^s z_k^s = f_k^s$, and $f_k^s z_l^s = 0$, $\forall\, l \neq k$. We then linearize by substituting $(z_k^s)^2 = z_k^s$, $\forall\, k$, $z_l^s z_k^s = 0$, $\forall\, l \neq k$ and by defining $v_{ik}^s \equiv u_i^s z_k^s$, $\forall\, i = 0, \ldots, n$, $k = 1, \ldots, K^s$ and $w_{ijk}^s \equiv x_{ij} z_k^s$, $\forall\,(i,j) \in \mathcal{A}$, $k = 1, \ldots, K^s$. After decomposition and simplification steps (see Appendix A), we obtain the following subproblem.

SP-LS$^s(\hat{x})$-RLT:

$$\min \sum_{k=1}^{K^s} f_k^s \qquad \text{Duals}$$

$$\begin{aligned}
\text{subject to: } & -m_k^s v_{nk}^s - b_k^s z_k^s + f_k^s \ge 0 \quad \forall k = 1, \ldots, K^s && A_k\\
& v_{nk}^s - \tau_k^s z_k^s \ge 0 \quad \forall k = 1, \ldots, K^s && B_k^+\\
& -v_{nk}^s + \tau_{k+1}^s z_k^s \ge 0 \quad \forall k = 1, \ldots, K^s && B_k^-\\
& v_{jk}^s - v_{ik}^s - d_{ij}^s z_k^s + (d_{ij}^s - g_{ij}^s) w_{ijk}^s \ge 0 \quad \forall (i,j) \in \mathcal{A},\ k = 1, \ldots, K^s && C_{ijk}\\
& -\sum_{k=1}^{K^s} w_{ijk}^s = -\hat{x}_{ij} \quad \forall (i,j) \in \mathcal{A} && D_{ij}\\
& -w_{ijk}^s + z_k^s \ge 0 \quad \forall (i,j) \in \mathcal{A},\ k = 1, \ldots, K^s && E_{ijk}\\
& \sum_{k=1}^{K^s} z_k^s = 1 && F\\
& w_{ijk}^s \ge 0 \quad \forall (i,j) \in \mathcal{A},\ k = 1, \ldots, K^s,\\
& z_k^s \ge 0 \quad \forall k = 1, \ldots, K^s.
\end{aligned}$$

Consider an optimal primal solution having objective function value $\hat{f}^s$ and completion time $\hat{u}_n^s$, such that $\hat{u}_n^s$ belongs to interval $k' \in \{1, \ldots, K^s\}$ (noting that $\hat{f}^s = m_{k'}^s \hat{u}_n^s + b_{k'}^s$). Define "slopes" ρ_k^L and ρ_k^R associated with interval k, where $\rho_k^L = (\hat{f}^s - (m_k^s \tau_k^s + b_k^s))/(\hat{u}_n^s - \tau_k^s)$ and $\rho_k^R = (\hat{f}^s - (m_k^s \tau_{k+1}^s + b_k^s))/(\hat{u}_n^s - \tau_{k+1}^s)$, $\forall k = 1, \ldots, K^s$. That is, ρ_k^L is the slope of the penalty from its current value to the value of the penalty function on the left interval value of segment k, and ρ_k^R is similarly defined for the right interval value of segment k. If $\hat{u}_n^s = \tau_{k'+1}^s$, we use the conventions $\rho_{k'}^R = m_{k'}^s$ and $\rho_{k'+1}^L = \infty$. Note that ρ_k^L and ρ_k^R are always nonnegative because all the penalty functions are nondecreasing. Define $\hat{\mathcal{A}}$ as a set containing all arcs in a critical path, and $\widehat{X}^b = \{(i,j) \in \hat{\mathcal{A}}: \hat{x}_{ij} = b\}$, for $b = 0$ and 1.

Proposition 1. *An optimal dual solution to* SP-LS$^s(\hat{x})$-RLT *is given as follows*:
- $A_k = 1$, $\forall k = 1, \ldots, K^s$.
- $C_{ijk} = 0$, $\forall (i,j) \notin \hat{\mathcal{A}}$, $k = 1, \ldots, K^s$.
- For all $k = 1, \ldots, k'$, $C_{ijk} = \max\{\rho_k^L, \rho_k^R, m_{k'}^s\}$, and for all $k = k'+1, \ldots, K^s$, $C_{ijk} = \min\{\rho_k^L, \rho_k^R, m_{k'}^s\}$, $\forall (i,j) \in \hat{\mathcal{A}}$. If $C_{ijk} \ge m_k^s$, $B_k^+ = 0$, $B_k^- = C_{ijk} - m_k^s$; otherwise, $B_k^- = 0$, $B_k^+ = m_k^s - C_{ijk}$.
- For arc $(i,j) \in \widehat{X}^1$, $D_{ij} = \min_{k=k', \ldots, K^s}\{(d_{ij}^s - g_{ij}^s) C_{ijk}\}$, and for arc $(i,j) \in \widehat{X}^0$, $D_{ij} = \max_{k=1, \ldots, k'-1}\{(d_{ij}^s - g_{ij}^s) C_{ijk}\}$.
- For all $k = 1, \ldots, K^s$, if arc $(i,j) \in \widehat{X}^1$, $E_{ijk} = (d_{ij}^s - g_{ij}^s) C_{ijk} - D_{ij}$, and if $(i,j) \in \widehat{X}^0$, $E_{ijk} = 0$.
- $F = \hat{f}^s + \sum_{(i,j) \in \widehat{X}^1} D_{ij}$.

Proof. See Appendix B. □

If η_s is obtained by solving the first-stage master problem for some $s \in \Omega$, such that $\eta_s < \hat{f}^s$, a Benders cut can be generated as

$$\eta_s \ge \hat{f}^s + \sum_{(i,j) \in \widehat{X}^1} D_{ij}(1 - x_{ij}) - \sum_{(i,j) \in \widehat{X}^0} D_{ij} x_{ij}. \tag{18}$$

Remark 1. Observe that we can compute the optimal dual values in Proposition 1 based on the current critical path length for the given scenario and its corresponding penalty. Hence, we generate Benders cuts in each scenario via the solution of critical path problems and avoid the direct solution of SP-LS$^s(\hat{x})$-RLT. This efficient cut-generation scheme is critical in reducing computational effort as evident in the computational results of §4. Also, note that the dual D_{ij} for arc $(i,j) \in \widehat{X}^1$ can be interpreted as an underestimate of the rate at which the penalty function would increase due to uninsuring arc (i,j), whereas D_{ij} for $(i,j) \in \widehat{X}^0$ is an overestimate of the rate of penalty reduction due to insuring arc (i,j).

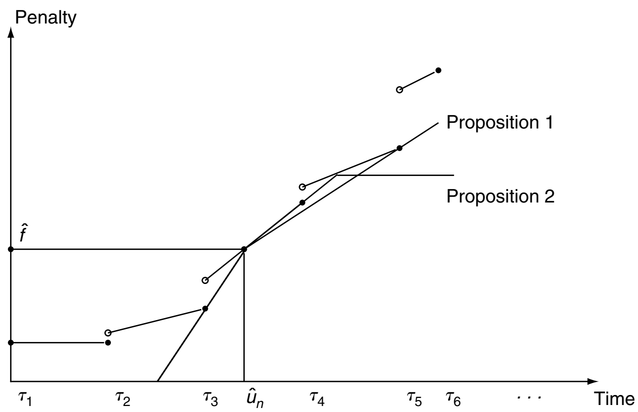
In fact, there exist several alternative optimal dual solutions to SP-LS$^s(\hat{x})$-RLT, which can yield different cutting planes for MP-PW, as shown in Proposition 2.

Proposition 2. *Denoting* $\Delta^s(\hat{x}) = (\tau_{k'+1}^s m_{k'+1}^s + b_{k'+1}^s - \hat{f}^s)$, *an alternative optimal dual solution to* SP-LS$^s(\hat{x})$-RLT *is given by modifying the dual values in Proposition* 1 *as follows*:
- For all $k = k'+1, \ldots, K^s$, set $C_{ijk} = 0$, $\forall (i,j) \in \hat{\mathcal{A}}$ and $B_k^- = 0$, $B_k^+ = m_k^s$.
- Arbitrarily order the arcs $(i,j) \in \widehat{X}^1$ and index them as $H = ((i,j)_1, \ldots, (i,j)_{|\widehat{X}^1|})$. Associating dual D_{ij_h} with arc $(i,j)_h \in \widehat{X}^1$, set $D_{(ij)_h} = \min\{(d_{(ij)_h}^s - g_{(ij)_h}^s) m_{k'}^s, \Delta^s(\hat{x}) - \sum_{l=1}^{h-1} D_{(ij)_l}\}$, $h = 1, \ldots, H$ and $E_{ijk} = (d_{ij}^s - g_{ij}^s) \cdot C_{ijk} - D_{ij}$, $\forall k = 1, \ldots, k'$, $E_{ijk} = 0$, $\forall k = k'+1, \ldots, K^s$.

Proof. The proof is similar to the proof of Proposition 1 (see Appendix C). □

We compare cutting planes (18) generated based on the duals from Propositions 1 and 2 in Figure 1. Note that both Propositions 1 and 2 essentially promise a decrease of $(d_{ij}^s - g_{ij}^s) q^L$ in η_s due to setting $x_{ij} = 1$ for $(i,j) \in \widehat{X}^0$, where q^L is the minimum slope of a tangent that underestimates Θ^s from τ_1^s to $\hat{u}_n^s$ and touches Θ^s at $u_n^s = \hat{u}_n^s$. However, duals generated from Proposition 1 will increase η_s by $(d_{ij}^s - g_{ij}^s) q^R$ due to setting $x_{ij} = 0$ for $(i,j) \in \widehat{X}^1$, where q^R is the maximum slope of a tangent that underestimates Θ^s from $\hat{u}_n^s$ to $\tau_{K^s+1}^s$ and touches Θ^s at $u_n^s = \hat{u}_n^s$. By contrast, Proposition 2 increases η^s by $(d_{ij}^s - g_{ij}^s) m_{k'}^s$ due to setting $x_{ij} = 0$ for $(i,j) \in \widehat{X}^1$, but only enforces the total increase in η^s up to a maximum of $\Delta^s(\hat{x})$. After this limit is reached, the rate $m_{k'}^s$ of increase in η^s is no longer necessarily valid. Duals set according to Proposition 2 limit the total

Figure 1 Comparison of Cuts Generated Based on Propositions 1 and 2



increase of η_s by $\Delta^s(\hat{x})$ by enforcing that $\sum_{(i,j)\in\widehat{X}^1} D_{ij} \le \Delta^s(\hat{x})$, and by greedily assigning positive values to D_{ij} in the order prescribed by H, until a total weight of $\Delta^s(\hat{x})$ has been assigned to those values, or until all $D_{ij} = (d^s_{ij} - g^s_{ij})m^s_{k'}$. For instance, suppose that $\widehat{X}^1 = \{(1,2),(2,3),(3,4)\}$, with $m^s_{k'} = 1$ and $(d^s_{12} - g^s_{12}) = 15$, $(d^s_{23} - g^s_{23}) = 13$, and $(d^s_{34} - g^s_{34}) = 10$. Then if $\Delta^s(\hat{x}) = 24$, we could obtain any of the following five sets of dual D_{ij} values for $(i,j)\in\widehat{X}^1$: $(D_{12}, D_{23}, D_{34}) = (15, 9, 0), (15, 0, 9), (11, 13, 0), (1, 13, 10)$, or $(14, 0, 10)$ by different permutations of H. In general, we could generate an exponential number of alternative optimal dual solutions according to Proposition 2, none of which dominates the other. Furthermore, many of these inequalities may have $D_{ij} = 0$ for several $(i,j)\in\widehat{X}^1$, particularly for those (i,j) that appear late in the ordering H. Alternatively, if $m^s_{k'}$ and/or $\Delta^s(\hat{x})$ are relatively large, then we can essentially scale the duals D_{ij} for $(i,j)\in\widehat{X}^1$ to ensure that duals D_{ij}, $(i,j)\in\widehat{X}^1$ set according to Proposition 2 would not be forced to zero (unless $(d^s_{ij} - g^s_{ij})m^s_{k'} = 0$). Note that no such scaling is necessary if $\Delta^s(\hat{x}) \ge \sum_{(i,j)\in\widehat{X}^1} (d^s_{ij} - g^s_{ij})m^s_{k'}$, because in that case every permutation H would yield the same set of dual values, in which $D_{ij} = (d^s_{ij} - g^s_{ij})m^s_{k'}$, $\forall (i,j)\in\widehat{X}^1$. Also, if $\Delta^s(\hat{x}) = 0$, scaling would not be possible. Otherwise, if $0 < \Delta^s(\hat{x}) \le \sum_{(i,j)\in\widehat{X}^1} (d^s_{ij} - g^s_{ij})m^s_{k'}$, then setting $D_{ij} = (d^s_{ij} - g^s_{ij})m^s_{k'}[\Delta^s(\hat{x})/\sum_{(i,j)\in\widehat{X}^1} (d^s_{ij} - g^s_{ij})m^s_{k'}]$ for all $(i,j)\in\widehat{X}^1$ yields the following inequality:

$$\eta_s \ge \hat{f}^s - \sum_{(i,j)\in\widehat{X}^0} D_{ij}x_{ij} + \frac{\Delta^s(\hat{x})}{\sum_{(i,j)\in\widehat{X}^1} (d^s_{ij} - g^s_{ij})} \cdot \sum_{(i,j)\in\widehat{X}^1} (d^s_{ij} - g^s_{ij})(1 - x_{ij}). \tag{19}$$

Remark 2. If $m^s_{k'} > q^R$, then according to Proposition 2, we greedily increase the coefficients associated with some arcs $(i,j)\in\widehat{X}^1$, while ensuring that $\sum_{(i,j)\in\widehat{X}^1} D_{ij}$ does not exceed $\Delta^s(\hat{x})$. In fact, we may generate an alternative cut of the form (18) by using any slope value $\bar{q}^R \in [q^R, m^s_{k'}]$, in which we set $D_{(ij)_h} = \min\{(d^s_{(ij)_h} - g^s_{(ij)_h})\bar{q}^R, \Delta^s(\hat{x}) - \sum_{l=1}^{h-1} D_{(ij)_l}\}$, $h = 1, \ldots, H$. Let this affine function be described by $\bar{q}^R u^s_n + \bar{b}$.

Note that when $\bar{q}^R > q^R$, there exists a first "crossing point" $CP \ge \hat{u}^s_n$, such that there exists an $\varepsilon > 0$, where $\Theta^s(CP + \delta) < \bar{q}^R(CP + \delta) + \bar{b}$, $\forall \delta \in (0, \varepsilon)$. We can underestimate Θ^s in this case by using a continuous three-segment piecewise-linear function: one from τ^s_1 to $\hat{u}^s_n$ having slope q^L (and intersecting Θ^s at $u^s_n = \hat{u}^s_n$), one from $\hat{u}^s_n$ to CP having slope $\hat{q}^R$, and the last from CP to $\tau^s_{K^s+1}$ having slope 0. Because the penalty function is nondecreasing, we ensure that this generated function underestimates Θ^s. We can then generate cuts of the form (18) based on an application of Proposition 2 to this three-segment underestimating function.

3.2. General Nonconvex Penalty Function

In this section, we extend our cutting-plane algorithms to more general nonconvex penalty functions. We now assume only that Θ^s is lower semicontinuous, nondecreasing and does not have infinite derivatives. Our approach essentially and dynamically approximates the penalty function using linear or two-segment concave piecewise-linear functions, based on the current critical path length and penalty function shape. We describe the details as follows:

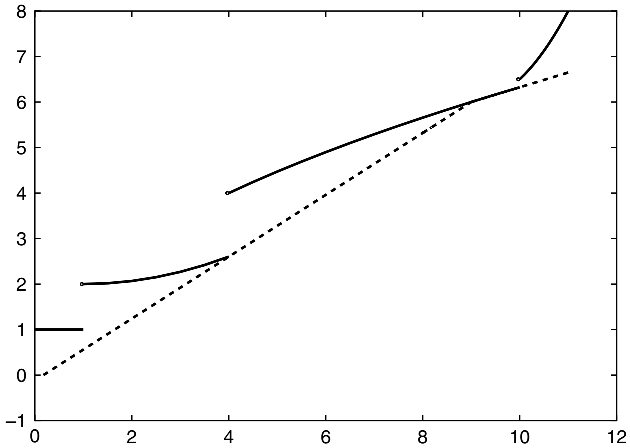
Step 1. Initialize the algorithm by setting $UB = \infty$ and by formulating MP-PW with no Benders inequalities.

Step 2. Solve the master problem MP-PW and obtain first-stage solution $\hat{x}$ and lower bound value LB. Solve $\text{CPM}^s(\hat{x})$ and obtain a critical path length $\hat{u}^s_n$ for each $s\in\Omega$, which yields an upper bound of $c\hat{x} + \sum_{s\in\Omega} e^s\Theta^s(\hat{u}^s_n)$ on the optimal objective value. If UB is larger than this upper bound value, then set $UB = c\hat{x} + \sum_{s\in\Omega} e^s\Theta^s(\hat{u}^s_n)$. If $LB = UB$, then terminate with optimal solution $\hat{x}$; else, continue to the next step.

Step 3. If $\eta_s < \Theta^s(\hat{u}^s_n)$ for some $s\in\Omega$, we can temporarily instate a piecewise-linear function that underestimates Θ^s. If a linear function $qu^s_n + q_0$ can underestimate Θ^s with $q\hat{u}^s_n + q_0 = \Theta^s(\hat{u}^s_n)$, then we generate a cut of the form (5). Otherwise, we underestimate Θ^s on the interval $[\tau^s_1, \hat{u}^s_n]$ with a linear function $q^L u^s_n + q^L_0$ and also on the interval $[\hat{u}^s_n, \tau^s_{K^s+1}]$ with a linear function $q^R u^s_n + q^R_0$, with $q^L > q^R$ and $q^L\hat{u}^s_n + q^L_0 = q^R\hat{u}^s_n + q^R_0 = \Theta^s(\hat{u}^s_n)$. We can then compute coefficients of (18) according to Propositions 1 or 2. We add cut (18) to MP-PW, and return to Step 2.

Note that inequality (18) generated for scenario s exactly approximates $\Theta^s(\hat{u}^s_n)$ at $\hat{x}$. Because all x variables are binary, there are a finite number of solutions to MP-PW, which ensures that this algorithm finitely reaches an optimal solution.

Figure 2 Illustration of Cutting-Plane Algorithm for a Nonconvex Penalty Function



Example 1. Suppose that the penalty function of scenario s is of the form

$$\Theta^s(t) = \begin{cases} 1 & 0 \le t \le 1, \\ 2 + (t-1)^2/15 & 1 < t \le 4, \\ 2\sqrt{t} & 4 < t \le 10, \\ 6 + (t-9)^2/2 & 10 < t \le 11. \end{cases}$$

Given a critical path length $\hat{u}_n^s = 9$, a penalty $\Theta^s(9) = 6$ is incurred. We underestimate Θ^s on the interval $[0, 9]$ by passing an affine function through $\hat{u}_n^s = 9$, $\Theta^s(\hat{u}_n^s) = 6$, with the smallest possible slope q^L that underestimates the function. Next, we repeat this procedure over the interval $[9, 11]$, obtaining a maximum slope q^R that underestimates Θ^s over this interval. We compute these slopes as follows (illustrated in Figure 2).

$$q^L = \frac{\hat{f}^s - \Theta^s(4)}{\psi^s(\hat{x}) - 4} = \frac{6 - 2.6}{9 - 4} = 0.68,$$

$$q^R = \frac{\hat{f}^s - \Theta^s(10)}{\psi^s(\hat{x}) - 10} = \frac{6 - 2\sqrt{10}}{9 - 10} = 2\sqrt{10} - 6.$$

If $\eta^s < 6$ in the solution of MP-PW, then we generate a Benders cut as

$$\eta_s \ge 6 + \sum_{(i,j)\in\hat{X}^1} (2\sqrt{10} - 6)(d_{ij}^s - g_{ij}^s)(1 - x_{ij}) - \sum_{(i,j)\in\hat{X}^0} 0.68(d_{ij}^s - g_{ij}^s)x_{ij}. \tag{20}$$

3.3. Chance-Constrained Problem

The previous analysis also permits us to consider a chance-constrained version of STIP as follows:

$$\mathbf{CC}_\epsilon\colon \quad \min\left\{ \sum_{(i,j)\in\mathcal{A}} c_{ij}x_{ij}\colon x \in X, \; \Pr\{\psi(x,\xi) \le \mathcal{T}\} \ge 1 - \epsilon \right\}, \tag{21}$$

where $X \subseteq B^{|\mathcal{A}|}$ forms a deterministic feasible region, ξ is a random vector with support $\xi \in \Xi \subset R^l$, and ψ: $R^n \times R^l \to R^m$ is a given constraint mapping that generates the critical path length given a first-stage decision x and ξ. Also, $\mathcal{T}$ is a random variable associated with the critical path length threshold, and ϵ is a risk level parameter chosen by the decision maker. Given a finite set of scenarios Ω, and ξ^s and $\mathcal{T}^s$ as the realization of ξ and $\mathcal{T}$ under scenario $s \in \Omega$, the chance constraint can be rewritten as

$$\sum_{s\in\Omega} \mathbb{I}(\psi(x,\xi^s) \le \mathcal{T}^s) \ge (1-\epsilon)|\Omega|, \tag{22}$$

where $\mathbb{I}(A)$ denotes whether event A is true (i.e., $\mathbb{I}(A) = 1$) or not (i.e., $\mathbb{I}(A) = 0$). Define variables $p^s \in \{0,1\}$, $\forall s \in \Omega$, such that $p^s = 1$ if a critical path length is permitted to violate the project target time $\mathcal{T}^s$ in scenario s, and $p^s = 0$ otherwise. Recalling that η_s is the sth subproblem objective, we formulate the master problem of CC_ϵ as follows:

$$\begin{aligned} \mathbf{CC}_\epsilon\textbf{-MP:} \quad & \min \sum_{(i,j)\in\mathcal{A}} c_{ij}x_{ij} \\ \text{subject to: } & m^{s,l}x + n^{s,l}\eta_s \ge o^{s,l} \quad \forall\, l \in L^s,\ s \in \Omega && (23) \\ & \sum_{s\in\Omega} \eta_s \le \lfloor \epsilon|\Omega| \rfloor && (24) \\ & x_{ij} \in \{0,1\} \quad \forall\,(i,j) \in \mathcal{A} \\ & \eta_s \ge 0 \quad \forall\, s \in \Omega, && (25) \end{aligned}$$

where $m^{s,l}$, $n^{s,l}$, and $o^{s,l}$ represent the coefficients associated with the lth Benders cut derived from the sth subproblem, which is formulated as

$$\begin{aligned} \mathbf{CC}_\epsilon\textbf{-SP}^s(\hat{x})\textbf{:} \quad & \min\ p^s \\ \text{subject to: } & u_j^s - u_i^s \ge d_{ij}^s - (d_{ij}^s - g_{ij}^s)\hat{x}_{ij} \quad \forall\,(i,j) \in \mathcal{A} && (26) \\ & u_n^s \le \mathcal{T}^s + Qp^s && (27) \\ & p^s \in \{0,1\}, && (28) \end{aligned}$$

where Q is once again an arbitrarily large constant. We multiply (26) and (27), as well as the bounds $x_{ij} \le 1$ and $x_{ij} \ge 0$, by p^s and $(1 - p^s)$, to generate the convex hull of solutions for which p^s is binary. We then linearize by substituting $(p^s)^2 = p^s$, $p^s(1-p^s) = 0$ and by defining $v_i^s \equiv u_i^s p^s$, $\forall\, i = 0, \ldots, n$, $w_{ij}^s \equiv x_{ij}p^s$, $\forall\,(i,j) \in \mathcal{A}$. After decomposition, the resulting subproblem is given by

$\mathbf{CC}_\epsilon\textbf{-SP}^s(\hat{x})\textbf{-RLT:}$

$$\begin{aligned} & \min\ p^s \\ \text{subject to: } & v_j^s - v_i^s + (d_{ij}^s - g_{ij}^s)w_{ij}^s - d_{ij}^s p^s \ge 0 \quad \forall\,(i,j) \in \mathcal{A} && (29) \\ & u_j^s - v_j^s - u_i^s + v_i^s - (d_{ij}^s - g_{ij}^s)w_{ij}^s + d_{ij}^s p^s \end{aligned}$$

$$\geq d_{ij}^s - (d_{ij}^s - g_{ij}^s)\hat{x}_{ij} \quad \forall (i,j) \in \mathcal{A} \tag{30}$$

$$-v_n^s + (\mathcal{T}^s + Q)p^s \geq 0 \tag{31}$$

$$-u_n^s + v_n^s - \mathcal{T}^s p^s \geq -\mathcal{T}^s \tag{32}$$

$$-w_{ij}^s \geq -\hat{x}_{ij} \quad \forall (i,j) \in \mathcal{A} \tag{33}$$

$$-w_{ij}^s + p^s \geq 0 \quad \forall (i,j) \in \mathcal{A} \tag{34}$$

$$w_{ij}^s - p^s \geq \hat{x}_{ij} - 1 \quad \forall (i,j) \in \mathcal{A} \tag{35}$$

$$w_{ij}^s \geq 0 \quad \forall (i,j) \in \mathcal{A}, \qquad p^s \geq 0. \tag{36}$$

After solving CC_ϵ-MP, if $\hat{\eta}_s = 1$, or if both $\hat{\eta}_s$ and the optimal objective value to $\text{CC}_\epsilon\text{-SP}^s(\hat{x})$-RLT equal zero, then no cut is generated. Otherwise, if $\hat{\eta}_s$ is less than the optimal objective to $\text{CC}_\epsilon\text{-SP}^s(\hat{x})$-RLT, we generate a Benders cut.

With respect to the structural constraints, we associate duals C_{ij}^-, C_{ij}^+, B^+, B^-, D_{ij}, E_{ij}^-, and E_{ij}^+ with constraints (29), (30), (31), (32), (33), (34), and (35), respectively. Recall that $\hat{\mathcal{A}}$ denotes a set containing all arcs in a critical path, and $\widehat{X}^b = \{(i,j) \in \hat{A}: \hat{x}_{ij} = b\}$, for $b = 0$ and 1.

Proposition 3. *Given $\hat{x}$ from the master problem, for each scenario s with $\hat{p}^s = 1$ (i.e., $\hat{u}_n^s > \mathcal{T}^s$), if $\hat{\eta}_s < \hat{p}^s = 1$, an optimal solution to* $\text{CC}_\epsilon\text{-SP}^s(\hat{x})$-RLT *is given as follows*:
- $C_{ij}^- = C_{ij}^+ = 1/(\hat{u}_n^s - \mathcal{T}^s)$, $\forall (i,j) \in \hat{\mathcal{A}}$.
- $C_{ij}^- = C_{ij}^+ = 0$, $\forall (i,j) \in \mathcal{A} \backslash \hat{\mathcal{A}}$.
- $B^+ = B^- = 1/(\hat{u}_n^s - \mathcal{T}^s)$.
- For all $(i,j) \in \widehat{X}^1$, $E_{ij}^+ = E_{ij}^- = (d_{ij}^s - g_{ij}^s)/(\hat{u}_n^s - \mathcal{T}^s)$, $D_{ij} = 0$.
- For all $(i,j) \in \widehat{X}^0$, $E_{ij}^+ = E_{ij}^- = D_{ij} = 0$.

Proof. The proof is similar to the proof of Proposition 1 (see Appendix D). □

For each scenario s with $\hat{p}^s = 1$, if $\hat{\eta}_s < \hat{p}^s$, a Benders cut is generated as

$$\begin{aligned}\eta_s &\geq \frac{\sum_{(i,j)\in\hat{\mathcal{A}}} d_{ij}^s - (d_{ij}^s - g_{ij}^s)x_{ij}}{\hat{u}_n^s - \mathcal{T}^s} - \frac{\mathcal{T}^s}{\hat{u}_n^s - \mathcal{T}^s} \\ &\quad + \sum_{(i,j)\in\widehat{X}^1} \frac{(d_{ij}^s - g_{ij}^s)}{\hat{u}_n^s - \mathcal{T}^s}(x_{ij} - 1) \\ &= \frac{\sum_{(i,j)\in\hat{\mathcal{A}}} d_{ij}^s - \sum_{(i,j)\in\widehat{X}^1}(d_{ij}^s - g_{ij}^s) - \mathcal{T}^s}{\hat{u}_n^s - \mathcal{T}^s} \\ &\quad - \sum_{(i,j)\in\hat{\mathcal{A}}\backslash\widehat{X}^1} \frac{(d_{ij}^s - g_{ij}^s)}{\hat{u}_n^s - \mathcal{T}^s} x_{ij} \\ &= 1 - \sum_{(i,j)\in\widehat{X}^0} \frac{(d_{ij}^s - g_{ij}^s)}{\hat{u}_n^s - \mathcal{T}^s} x_{ij}.\end{aligned} \tag{37}$$

4. Computational Results

We demonstrate the computational efficacy of our cutting-plane algorithms for the expectation and chance-constrained problems by testing our algorithms on 15 randomly generated network instances. Table 1 provides the parameters used to generate the instances, where $|\mathcal{N}|$ gives the number of nodes and "Degree range" denotes the minimum and maximum degrees of each node allowed in the initial phase of graph generation. We generate each instance as a topologically ordered graph (i.e., where $(i,j) \in \mathcal{A}$ only if $i < j$) by implementing the following procedures. We begin by initializing $\mathcal{A} = \varnothing$, and then in a loop, we randomly generate two nodes $i, j \in \mathcal{N}$, where $i < j$, $(i,j) \notin \mathcal{A}$, and the degree of both i and j is strictly smaller than the maximum value of the degree range. We add arc (i,j) to the graph and increase the degrees of i and j by one. We repeat this procedure until the degrees of all nodes lie within the degree range. After this initial phase is complete, we ensure that there exists at least one path from node 0 to node i, and one from node i to node n, for all $i = 1, \ldots, n-1$. If not, we artificially construct such path(s) from node 0 to node i or from node i to node n, and make sure that each path contains at least $\min\{0.2|\mathcal{N}|, \lceil 0.5i \rceil\}$ nodes for paths connecting 0 to i, and $\min\{0.2|\mathcal{N}|, \lceil 0.5(|\mathcal{N}| - i) \rceil\}$ nodes for paths connecting i to n. (Note that we potentially violate the maximum value of the degree range at some nodes after adding these additional paths.) We generate five such instances for each combination of $|\mathcal{N}|$ and degree range, and we report the total number of arcs generated for each instance in Table 1.

Table 1 **Test Instances**

$\lvert\mathcal{N}\rvert$	Degree range	$\lvert\mathcal{A}\rvert$ for each instance 1	2	3	4	5
30	[5, 15]	332	322	348	388	394
50	[10, 20]	850	814	870	802	856
70	[15, 25]	1,640	1,606	1,638	1,454	1,490

For each arc $(i,j) \in \mathcal{A}$, we randomly generate a typical task duration value from a uniform distribution over the interval [10, 300]. To generate scenario data, we examine the practical case in which a task is more likely to be delayed than completed earlier and where the duration of delays exceeds the amount of time by which a task could be early. Hence, we generate d_{ij}^s in scenario s by multiplying the typical task duration for $(i,j) \in \mathcal{A}$ by a random value uniformly distributed over the interval [0.9, 1.5]. We randomly generate g_{ij}^s for arc $(i,j) \in \mathcal{A}$, scenario s, by generating g_{ij}^s from the uniformly distributed $[0.5d_{ij}^s, 0.7d_{ij}^s]$. The cost c_{ij} to insure arc $(i,j) \in \mathcal{A}$ is uniformly distributed over the interval [25, 50]. Finally, we round all the values of d_{ij}^s, g_{ij}^s, and c_{ij} to the nearest integer values.

4.1. Expectation-Based Penalty Function Cases

Our first experiment tests the computational efficacy of our procedures for the cases of convex penalty functions (§2) and nonconvex piecewise-linear lower semicontinuous penalty functions (§3.1). Here we employ the SAA method (see Kleywegt et al. 2001, Mak et al. 1999, Norkin et al. 1998, and Appendix E), an exterior sampling method designed to provide bounds on stochastic programs. In each case, we vary the sample size $N = |\Omega|$ to examine the trade-off in narrowing the gap between computed statistical upper and lower bounds and increasing solution time by using relatively large values of N. We test all instances with sample sizes of $N = 50$, 100, and 200 scenarios. We use three and five penalty function segments, denoted as "three-segment" and "five-segment," respectively, for both convex and nonconvex penalty cases. We name all instances as **w-x-y-z**, where $\mathbf{w} = $ C or Nc corresponding to convex and nonconvex cases, respectively, and **x**, **y**, **z** are the number of function segments, number of nodes, and instance number, respectively. For instance, Nc-5-30-3 is a nonconvex five-segment instance using the third 30-node network.

To generate the penalty functions, we first compute an original critical path length $\hat{u}_n^s$ according to the uninsured duration times d_{ij}^s. We compute the threshold values $\tau_1^s, \ldots, \tau_{K^s+1}^s$ by setting $\tau_i^s = (0.7 + 0.3(i-1)/K^s)\hat{u}_n^s,\ \forall\, i = 1, \ldots, K^s + 1$, so that $\tau_1^s = 0.7\hat{u}_n^s$, $\tau_{K^s+1}^s = \hat{u}_n^s$, and all intervals are evenly distributed. If completion time is within the target due time (i.e., less than τ_1^s), no penalty is incurred. The remainder of the penalty function is generated as follows.

For the convex penalty case, we first generate the left-hand-point function value $f_1^L = 0$ of segment 1, and then we generate an increasing series of slopes m_k^s of each segment k such that $0 < m_1^s < m_2^s < \cdots < m_{K^s}^s$ (we use $m_1^s = 1.2$, $m_k^s = 1.2m_{k-1}^s$, and $m_1^s = 1.1$, $m_k^s = 1.1m_{k-1}^s$ for the three-segment and five-segment cases, respectively, $\forall\, k = 2, \ldots, K^s$). We compute f_k^L at the left-hand-point of each segment k as $f_k^L = f_{k-1}^L + m_k^s(\tau_k - \tau_{k-1}),\ \forall\, k = 2, \ldots, K^s$, to ensure convexity of the penalty function. For the nonconvex case, f_1^L is uniformly distributed over the interval [100, 400]. We randomly generate the slopes m_k^s according to a uniform distribution over the interval [0.2, 0.9]. To ensure that the function is increasing, we set $f_k^L = 1.05(f_{k-1}^L + m_k^s(\tau_k^s - \tau_{k-1}^s))$, $\forall\, k = 2, \ldots, K^s$, and thus the increasing piecewise-linear function becomes discontinuous and nonconvex.

For each instance, we use $M = 20$ as the number of samples, and the size of the reference sample is set to $N' = 10{,}000$ scenarios. All decomposition algorithms are implemented using CPLEX 11.0 (ILOG 2008) via ILOG Concert Technology 2.5, and all computations are performed on a SUN UltraSpace-III with a 900 MHz processor and 2.0 GB RAM. Computational times are reported in CPU seconds. We allow a one-hour (3,600 seconds) time limit.

4.1.1. Computational Results of Expectation-Based Models. For the convex case, we use cut (5), and for the nonconvex case, we compute cut (18) using optimal dual values according to Proposition 1, and present the computational results in Tables 2–5. For these tables, $t_{\max}$, $t_{\min}$, and t_{avg} represent the maximum, minimum, and average CPU seconds for each instance over all $M = 20$ samples, respectively. If our algorithm fails to solve some sample within the time limit, we report "LIMIT" in $t_{\max}$ and present the average CPU time of all solvable samples in t_{avg}. Columns labeled "Iter" and "Cuts" represent the average number of times that we solve the master problem and the average number of Benders cuts (5) or (18) generated before achieving optimality, respectively. Columns labeled "LB" and "UB" denote the statistical lower bound and the best (minimum) upper bound of the optimal objective function value using the reference sample, respectively. The column labeled "Gap" represents the difference between LB and UB as a percentage of the lower bound.

Comparing the results with respect to the convex and nonconvex cases, the latter requires more iterations and cuts generation, and thus increases the CPU time. We observe that the optimality gaps improve by increasing the sample size, N, of scenarios. However, increasing N also leads to an increase in CPU times and in cutting-plane generation. For instance, using $N = 200$ scenarios, we reduce the optimality gaps associated with all 15 instances to less than 1% in all computational experiments represented by Tables 2–5. The average number of cuts generated by each sample is approximately a linear function of scenarios, and the average CPU time increases by more than five times for some instances compared to the case of $N = 100$. Indeed, some samples of instances having 50 and 70 nodes cannot be solved within the time limit for the five-segment nonconvex penalty case.

Remark 3. Because we only need to solve a critical path problem to obtain the Benders cutting planes required for our algorithm, we save significant computational effort compared to approaches that (a) directly solve the nondecomposed mixed-integer model or (b) decompose the model but explicitly solve subproblems SP-LS$^s(\hat{x})$-RLT by linear programming to obtain the duals (as opposed to our dual recovery procedure given in Proposition 1). To illustrate the computational importance of our approach, we solved instance Nc-3-30-1 with 50 scenarios by applying CPLEX to the nondecomposed mixed-integer model. Each of the $M = 20$ samples took at least six CPU hours to solve using this approach, compared with an average of 20.47 seconds using our decomposition methodology. Using the same instance, we also decomposed the problem

Table 2 Computational Results for Three-Segment Convex Instances

$N = \|\Omega\|$	Instances	t_{max}	t_{min}	t_{avg}	Iter	Cuts	LB	UB	Gap (%)
$N = 50$	C-3-30-1	0.64	0.10	0.30	5.30	71.20	234.44	240.18	2.45
	C-3-30-2	0.65	0.07	0.26	4.50	58.10	44.81	47.20	5.33
	C-3-30-3	1.11	0.19	0.47	5.40	81.30	290.37	295.41	1.74
	C-3-30-4	1.13	0.19	0.59	5.65	72.80	285.46	289.24	1.32
	C-3-30-5	1.24	0.18	0.62	6.10	98.45	111.87	114.87	2.68
	C-3-50-1	3.28	0.27	1.14	5.30	74.20	167.22	172.14	2.94
	C-3-50-2	2.70	0.17	0.99	4.85	58.10	57.15	59.10	3.41
	C-3-50-3	3.85	0.29	1.56	5.05	95.05	188.03	191.18	1.68
	C-3-50-4	2.42	0.22	0.94	4.25	56.10	94.79	96.15	1.43
	C-3-50-5	3.21	0.25	1.16	5.15	76.30	170.37	174.90	2.66
	C-3-70-1	5.53	0.99	2.63	5.20	97.40	352.36	356.34	1.13
	C-3-70-2	5.45	1.00	2.52	5.30	107.60	311.12	319.69	2.75
	C-3-70-3	5.59	0.96	2.62	5.30	114.15	157.63	158.12	0.31
	C-3-70-4	4.47	0.84	2.03	4.95	72.50	239.35	245.89	2.73
	C-3-70-5	4.56	0.94	2.30	5.10	80.05	193.19	196.00	1.45
$N = 100$	C-3-30-1	1.25	0.23	0.55	4.95	139.75	238.89	242.07	1.33
	C-3-30-2	1.27	0.11	0.52	4.60	124.20	46.05	47.13	2.35
	C-3-30-3	2.08	0.27	0.96	5.30	161.45	290.25	293.41	1.09
	C-3-30-4	2.35	0.31	1.14	5.45	146.60	290.08	290.95	0.30
	C-3-30-5	2.64	0.30	1.39	6.00	194.70	108.50	110.77	2.09
	C-3-50-1	6.48	0.42	1.98	5.35	141.10	162.23	166.07	2.37
	C-3-50-2	4.85	0.31	1.77	4.80	114.25	57.55	58.10	0.96
	C-3-50-3	7.26	0.39	2.52	5.00	169.55	190.37	191.18	0.43
	C-3-50-4	4.46	0.31	1.68	4.05	110.65	93.56	94.54	1.05
	C-3-50-5	6.04	0.41	1.97	5.30	148.40	173.49	174.90	0.81
	C-3-70-1	15.40	2.25	6.86	5.25	181.65	353.47	356.34	0.81
	C-3-70-2	15.28	2.18	6.70	5.65	202.95	312.14	316.70	1.46
	C-3-70-3	17.01	2.23	7.86	5.35	217.50	155.06	156.97	1.23
	C-3-70-4	13.25	1.90	5.33	4.85	144.50	239.37	242.56	1.33
	C-3-70-5	14.23	1.97	6.11	5.05	155.00	193.33	194.47	0.59
$N = 200$	C-3-30-1	2.10	0.44	0.96	4.35	230.35	241.08	242.07	0.41
	C-3-30-2	3.08	0.24	1.12	4.25	241.25	46.81	47.13	0.68
	C-3-30-3	4.35	0.42	1.91	5.15	320.30	292.25	292.68	0.15
	C-3-30-4	5.06	0.42	1.94	5.40	292.60	290.72	290.95	0.08
	C-3-30-5	5.00	0.45	2.40	5.95	322.75	109.73	110.77	0.95
	C-3-50-1	12.47	0.65	4.74	5.20	263.35	162.34	163.12	0.48
	C-3-50-2	9.31	0.55	4.08	4.55	218.05	57.91	58.10	0.33
	C-3-50-3	13.42	0.60	4.91	5.10	322.45	190.26	190.72	0.24
	C-3-50-4	8.15	0.55	3.91	4.25	208.95	93.95	94.54	0.63
	C-3-50-5	11.89	0.73	4.81	5.15	279.50	174.52	174.63	0.06
	C-3-70-1	21.40	4.02	10.39	4.80	339.65	353.91	356.34	0.69
	C-3-70-2	23.03	4.23	9.60	5.15	397.55	314.77	316.70	0.61
	C-3-70-3	23.87	4.10	11.33	5.30	409.50	155.96	156.97	0.65
	C-3-70-4	17.52	3.28	7.54	4.85	280.35	241.21	242.56	0.56
	C-3-70-5	19.49	3.49	8.57	5.05	317.60	193.98	194.47	0.25

and solved the RLT-enhanced subproblem SP-LS$^s(\hat{x})$-RLT by linear programming rather than by our dual recovery technique, and none of the 20 samples were solved within the one-hour time limit.

4.1.2. Analysis of Insured Arc Characteristics. In this part, we provide insights pertaining to optimal solutions obtained by our expectation-based STIP models. Let $\bar{d}_{ij}$ and $\bar{g}_{ij}$, respectively, denote average uninsured and insured task durations of arc (i, j) over all scenarios; thus, $\bar{d}_{ij} - \bar{g}_{ij}$ represents the average duration-reduction value for arc (i, j) due to its insurance. Here, we examine the extent to which small ratios of an arc's cost-to-duration-reduction ratio $c_{ij}/(\bar{d}_{ij} - \bar{g}_{ij})$ influences whether or not the arc will be insured in the optimal solution we obtain. For a given instance, we order all arcs $(i, j) \in \mathcal{A}$ in nondecreasing order of their $c_{ij}/(\bar{d}_{ij} - \bar{g}_{ij})$ values, and we examine the frequency in which arcs at different portions of this spectrum are insured in the obtained optimal solution.

We conduct this experiment on all **w-5-30-z** and **w-5-50-z** instances. In Figure 3, the arcs are partitioned into groups such that the top 10% of arcs ordered as above belong to the first group (labeled "10%"), followed by the next top 10% of arcs in the second group (labeled "20%"), and so on. These groups are depicted

Table 3 Computational Results for Five-Segment Convex Instances

$N = \|\Omega\|$	Instances	t_{max}	t_{min}	t_{avg}	Iter	Cuts	LB	UB	Gap (%)
$N = 50$	C-5-30-1	1.29	0.23	0.59	5.20	70.05	267.40	274.41	2.62
	C-5-30-2	1.27	0.24	0.48	4.00	58.05	78.36	80.32	2.50
	C-5-30-3	1.54	0.32	0.64	5.25	75.15	317.95	324.35	2.01
	C-5-30-4	1.50	0.28	0.48	4.40	67.30	328.31	338.09	2.98
	C-5-30-5	1.86	0.31	0.94	6.00	82.75	158.84	163.09	2.68
	C-5-50-1	4.67	0.48	1.44	5.10	71.85	148.53	152.42	2.62
	C-5-50-2	5.09	0.30	1.05	4.35	56.35	73.35	75.06	2.33
	C-5-50-3	8.49	0.58	1.66	5.00	87.10	187.97	191.29	1.77
	C-5-50-4	4.37	0.42	1.22	4.50	61.95	134.84	138.46	2.68
	C-5-50-5	6.56	0.47	2.31	5.20	73.30	164.25	168.91	2.84
	C-5-70-1	9.40	1.53	3.84	5.05	97.15	354.36	361.41	1.99
	C-5-70-2	10.71	1.40	3.97	5.25	103.05	316.45	323.53	2.24
	C-5-70-3	11.49	1.55	3.86	5.40	112.00	191.31	191.20	−0.06
	C-5-70-4	7.90	1.00	3.03	4.90	70.15	232.47	238.96	2.79
	C-5-70-5	8.77	1.13	3.11	5.05	78.45	230.65	235.73	2.20
$N = 100$	C-5-30-1	2.69	0.29	1.43	4.95	142.50	266.95	270.98	1.51
	C-5-30-2	2.48	0.30	1.17	4.05	119.30	73.08	74.53	1.98
	C-5-30-3	2.80	0.41	1.28	5.35	142.55	321.47	324.35	0.90
	C-5-30-4	2.62	0.39	1.00	4.80	135.25	332.40	334.51	0.63
	C-5-30-5	3.34	0.49	2.07	6.00	169.65	155.75	157.96	1.42
	C-5-50-1	9.77	0.92	3.04	5.20	140.45	145.86	148.37	1.72
	C-5-50-2	9.16	0.53	2.49	4.40	112.30	73.97	75.06	1.47
	C-5-50-3	16.47	1.07	4.38	5.15	170.50	185.38	186.95	0.85
	C-5-50-4	8.74	0.63	2.06	4.65	117.30	134.25	136.14	1.41
	C-5-50-5	13.18	0.87	3.65	5.25	141.60	166.40	168.91	1.51
	C-5-70-1	19.63	2.81	7.34	5.20	184.65	356.45	361.41	1.39
	C-5-70-2	19.92	3.35	8.44	5.20	203.55	312.72	317.67	1.58
	C-5-70-3	23.92	2.84	7.91	5.45	216.85	190.81	191.20	0.20
	C-5-70-4	17.72	1.93	5.74	4.95	135.20	231.41	235.03	1.56
	C-5-70-5	18.25	2.00	5.88	5.10	156.70	231.37	235.73	1.88
$N = 200$	C-5-30-1	7.90	1.23	3.26	4.90	238.65	270.01	270.98	0.36
	C-5-30-2	6.04	0.77	3.09	4.25	238.15	73.19	73.35	0.22
	C-5-30-3	7.32	0.93	3.44	5.00	280.65	324.22	324.35	0.04
	C-5-30-4	6.97	0.89	3.11	4.95	268.35	334.08	334.51	0.13
	C-5-30-5	9.73	1.25	5.24	5.90	328.10	157.05	157.96	0.58
	C-5-50-1	26.10	1.87	8.82	5.10	273.50	148.09	148.37	0.19
	C-5-50-2	24.01	1.65	7.32	4.35	214.65	73.39	73.88	0.67
	C-5-50-3	41.48	1.92	9.32	4.95	331.60	185.00	185.16	0.09
	C-5-50-4	23.23	1.25	7.34	4.45	212.75	134.96	135.24	0.21
	C-5-50-5	37.01	2.57	10.12	5.05	269.10	168.70	168.91	0.12
	C-5-70-1	82.86	7.95	21.40	4.95	346.60	359.40	360.09	0.19
	C-5-70-2	79.58	8.78	24.15	5.00	389.85	311.58	314.50	0.94
	C-5-70-3	95.13	8.29	22.98	5.25	407.95	190.94	191.20	0.14
	C-5-70-4	59.26	6.58	17.42	4.75	256.10	231.95	232.86	0.39
	C-5-70-5	63.72	6.21	18.29	5.00	286.70	231.72	233.24	0.66

on the horizontal axis, and the vertical axis represents the percentage of arcs in each of the 10 groups that are insured in the optimal solution obtained. We see that arcs (i, j) having very high values of $c_{ij}/(\bar{d}_{ij} - \bar{g}_{ij})$ relative to other arcs' ratios are not likely to be insured. Indeed, no arcs in the upper "20%" of cost-to-duration-ratio were insured in optimal solutions to any of the instances tested here. However, no trend is evident regarding which of the remaining arcs will be selected in an optimal solution. This underscores the difficulty of the problem and the necessity of using sophisticated approaches for their solution.

4.1.3. Analysis of the Persistency of the First-Stage Optimal Solution. In this part, we test the notion that one may be able to anticipate which arcs will be insured at optimality by solving a series of deterministic task-insurance instances, one corresponding to each possible scenario. Specifically, for each scenario $s \in \Omega$, we could solve a deterministic problem as $\min\{cx + f(\psi(x, \xi^s)): x \in \{0, 1\}^{|\mathcal{A}|}\}$ and obtain its optimal first-stage solution as $x^*(\xi^s)$. For each $(i, j) \in A$, we then compute the percentage of these $|\Omega|$ instances in which (i, j) is insured (i.e., given by $(\sum_{s\in\Omega} x^*_{ij}(\xi^s))/|\Omega|$). The arcs that are insured with high frequency are said to be persistent. A closely

Table 4 Computational Results for Three-Segment Nonconvex Instances

$N = \|\Omega\|$	Instances	t_{max}	t_{min}	t_{avg}	Iter	Cuts	LB	UB	Gap (%)
$N = 50$	Nc-3-30-1	95.14	4.02	20.47	30.10	1,229.40	514.66	518.32	0.71
	Nc-3-30-2	82.59	2.83	17.55	27.55	1,083.50	312.70	317.21	1.44
	Nc-3-30-3	120.42	6.43	27.68	34.20	1,399.90	569.97	569.70	−0.05
	Nc-3-30-4	153.03	6.27	32.88	35.40	1,406.70	553.62	559.21	1.01
	Nc-3-30-5	172.30	7.91	35.62	34.75	1,483.55	378.65	383.32	1.23
	Nc-3-50-1	200.23	8.65	37.80	29.35	1,237.45	436.45	441.34	1.12
	Nc-3-50-2	154.37	4.82	30.45	27.10	1,006.20	323.29	326.95	1.13
	Nc-3-50-3	260.59	9.14	43.68	34.55	1,348.25	436.31	441.64	1.22
	Nc-3-50-4	152.45	5.03	29.87	26.90	1,012.80	400.65	406.00	1.34
	Nc-3-50-5	229.80	8.98	40.57	31.45	1,246.95	410.17	412.81	0.64
	Nc-3-70-1	87.43	19.48	46.33	33.55	1,263.60	615.83	621.64	0.94
	Nc-3-70-2	89.37	17.25	44.79	34.80	1,309.45	589.02	601.37	2.10
	Nc-3-70-3	82.56	14.31	41.11	33.15	1,254.35	439.15	445.72	1.50
	Nc-3-70-4	77.83	13.96	32.87	30.65	1,196.75	481.99	490.46	1.76
	Nc-3-70-5	74.70	10.23	35.64	30.90	1,188.50	480.51	486.85	1.32
$N = 100$	Nc-3-30-1	254.35	11.89	58.99	29.55	2,394.45	516.87	518.32	0.28
	Nc-3-30-2	222.49	10.21	34.83	27.40	2,090.20	314.56	316.89	0.74
	Nc-3-30-3	370.88	15.89	69.56	33.50	2,553.45	568.66	569.70	0.18
	Nc-3-30-4	346.90	14.37	79.18	34.95	2,798.05	557.81	559.21	0.25
	Nc-3-30-5	403.64	17.65	87.68	33.85	2,800.15	381.49	383.01	0.40
	Nc-3-50-1	437.24	20.67	74.26	29.20	2,321.40	432.50	436.01	0.81
	Nc-3-50-2	342.95	14.30	62.48	26.50	1,999.60	321.64	323.83	0.68
	Nc-3-50-3	528.47	21.52	83.20	33.25	2,596.70	435.42	439.67	0.98
	Nc-3-50-4	326.00	14.21	59.61	27.15	2,007.85	400.12	403.60	0.87
	Nc-3-50-5	495.38	19.73	81.77	30.80	2,485.65	410.56	412.81	0.55
	Nc-3-70-1	306.38	34.62	167.82	32.60	2,410.90	617.29	619.21	0.31
	Nc-3-70-2	357.65	37.28	178.45	33.75	2,568.30	587.26	591.28	0.68
	Nc-3-70-3	317.68	35.91	157.18	32.85	2,419.35	441.68	445.72	0.91
	Nc-3-70-4	299.14	25.04	139.21	30.20	2,398.70	485.12	490.46	1.10
	Nc-3-70-5	276.53	26.73	133.69	31.05	2,359.05	479.55	483.04	0.73
$N = 200$	Nc-3-30-1	942.50	22.80	206.10	30.35	5,024.85	516.19	516.28	0.02
	Nc-3-30-2	852.00	28.69	90.62	26.90	4,043.20	314.87	315.02	0.05
	Nc-3-30-3	999.57	20.17	200.14	34.00	4,977.50	568.29	568.36	0.01
	Nc-3-30-4	1,043.68	29.22	235.53	33.95	5,431.10	558.04	558.28	0.04
	Nc-3-30-5	1,125.41	36.35	289.64	34.30	5,596.55	380.25	380.57	0.08
	Nc-3-50-1	1,192.49	61.24	200.12	30.00	4,500.80	434.50	435.24	0.17
	Nc-3-50-2	967.67	43.37	177.95	26.85	4,040.15	323.23	323.83	0.19
	Nc-3-50-3	1,521.38	59.41	239.02	32.65	4,901.40	437.28	437.56	0.06
	Nc-3-50-4	910.55	37.69	169.34	27.55	3,960.35	402.79	403.60	0.20
	Nc-3-50-5	1,327.53	53.20	212.68	31.20	4,652.65	412.59	412.81	0.05
	Nc-3-70-1	1,278.52	98.42	459.77	33.10	4,733.85	619.09	621.64	0.41
	Nc-3-70-2	1,430.22	110.27	490.36	33.95	4,902.15	587.63	591.28	0.62
	Nc-3-70-3	1,350.06	106.52	445.80	32.35	4,782.45	443.17	443.95	0.18
	Nc-3-70-4	1,101.30	77.78	401.22	30.00	4,459.25	487.84	490.46	0.54
	Nc-3-70-5	1,062.79	81.93	392.76	30.75	4,507.30	481.28	483.04	0.37

related study was published by Bertsimas et al. (2006) for computing the persistency of binary variables (i.e., the probability that the variable will equal one at optimality) in discrete optimization problems under objective uncertainty with only partial information on the distribution of the objective coefficients. Here we empirically investigate whether persistent arcs correspond to those that are insured in the optimal STIP solution.

We test this hypothesis on instances Nc-3-30-1 and Nc-5-30-1 with a sample size of $N = 200$ and present the results in Table 6. The top row, labeled "Arc no.," gives the labels of arcs that were insured in at least one deterministic task-insurance instance (i.e., $x^*_{ij}(\xi^s) = 1$ for some $s \in \Omega$). The rows for Nc-3-30-1 and Nc-5-30-1 state the number of times that each arc appears in a deterministic task-insurance solution (out of 200 scenarios). For instance, for Nc-3-30-1, there are 50 scenarios in which arc 5 is insured out of the 200 deterministic task-insurance instances, and for Nc-5-30-1, there are 53 such scenarios in which arc 5 is insured. The arcs insured in the (unique, in both cases) optimal STIP solution we obtain are marked with a superior "a" in each row. (Arcs not depicted in Table 6 were not insured in the optimal STIP solution.)

Observe that optimality of persistent arcs does not hold in general, in the sense that arcs insured in a

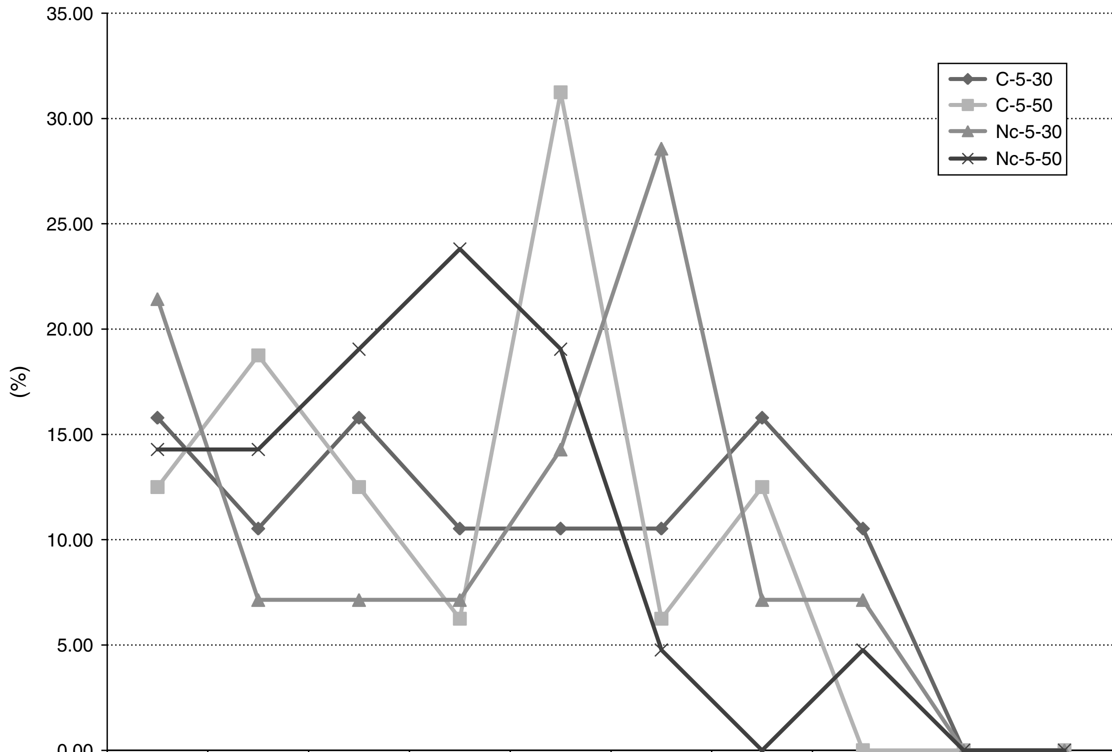
Table 5 Computational Results for Five-Segment Nonconvex Instances

$N = \|\Omega\|$	Instances	t_{max}	t_{min}	t_{avg}	Iter	Cuts	LB	UB	Gap (%)
$N = 50$	Nc-5-30-1	276.29	4.90	61.62	38.85	1,646.05	520.26	525.01	0.91
	Nc-5-30-2	222.87	3.76	53.27	30.25	1,521.35	316.33	319.62	1.04
	Nc-5-30-3	298.64	5.12	68.36	39.35	1,788.50	570.43	575.90	0.96
	Nc-5-30-4	336.78	6.22	82.19	38.10	1,893.05	559.66	563.63	0.71
	Nc-5-30-5	359.63	7.03	88.74	40.35	2,001.50	382.94	386.12	0.83
	Nc-5-50-1	594.37	8.28	178.65	42.65	1,865.25	438.92	443.91	1.14
	Nc-5-50-2	637.69	5.04	100.30	34.30	1,675.40	321.45	325.64	1.30
	Nc-5-50-3	787.42	9.54	189.61	45.75	2,011.50	442.37	446.88	1.02
	Nc-5-50-4	435.62	5.56	93.17	43.00	1,724.65	404.68	409.53	1.20
	Nc-5-50-5	677.34	8.95	143.55	43.70	2,189.75	415.80	421.01	1.25
	Nc-5-70-1	1,009.37	20.96	209.77	54.30	2,219.55	620.17	625.58	0.87
	Nc-5-70-2	1,191.64	23.58	219.47	56.65	2,403.85	592.54	609.01	2.78
	Nc-5-70-3	1,006.52	21.04	203.61	53.30	2,283.35	479.38	485.27	1.23
	Nc-5-70-4	978.43	18.72	184.65	50.55	2,154.40	475.24	483.33	1.70
	Nc-5-70-5	925.80	19.25	176.19	51.65	2,199.15	480.28	482.87	0.54
$N = 100$	Nc-5-30-1	561.99	12.63	158.95	39.80	3,250.20	517.42	523.42	1.16
	Nc-5-30-2	453.23	7.29	100.11	32.10	3,008.65	318.21	319.29	0.34
	Nc-5-30-3	594.84	10.11	123.90	39.10	3,506.90	574.85	575.90	0.18
	Nc-5-30-4	627.81	11.30	156.47	40.05	3,582.95	563.89	565.31	0.25
	Nc-5-30-5	700.55	13.98	162.44	41.15	3,996.85	385.81	386.58	0.20
	Nc-5-50-1	1,236.28	22.91	394.87	43.10	3,788.25	437.97	441.62	0.83
	Nc-5-50-2	1,485.67	16.39	269.48	34.00	3,349.60	322.96	325.64	0.83
	Nc-5-50-3	1,732.10	25.88	413.26	45.65	4,211.50	444.34	446.88	0.57
	Nc-5-50-4	1,323.14	15.64	221.59	41.35	3,405.60	402.38	405.15	0.69
	Nc-5-50-5	1,568.62	24.30	347.63	42.10	4,300.20	417.17	419.73	0.61
	Nc-5-70-1	LIMIT	237.45	1,617.25	53.30	4,183.25	621.85	625.58	0.60
	Nc-5-70-2	LIMIT	272.37	1,638.91	56.70	4,297.60	597.88	605.35	1.25
	Nc-5-70-3	LIMIT	214.58	1,594.32	53.90	4,164.30	477.12	481.23	0.86
	Nc-5-70-4	2,947.81	190.46	1,323.96	50.20	4,005.95	475.98	483.33	1.54
	Nc-5-70-5	3,245.89	192.31	1,345.88	51.80	4,017.30	477.84	480.29	0.51
$N = 200$	Nc-5-30-1	1,629.64	49.53	394.75	35.50	6,024.40	523.31	523.42	0.02
	Nc-5-30-2	1,299.30	27.41	293.64	30.15	5,996.70	318.80	319.29	0.15
	Nc-5-30-3	1,701.52	31.27	360.11	36.35	6,893.60	573.84	574.55	0.12
	Nc-5-30-4	1,805.13	30.06	407.82	37.45	7,018.25	564.13	564.37	0.04
	Nc-5-30-5	1,950.68	33.28	459.35	38.70	7,845.80	384.56	384.88	0.08
	Nc-5-50-1	3,468.23	81.24	874.52	41.95	7,412.65	437.65	440.21	0.58
	Nc-5-50-2	LIMIT	63.59	637.95	33.65	6,700.15	323.00	323.93	0.29
	Nc-5-50-3	LIMIT	96.72	903.21	44.85	8,395.30	446.47	446.88	0.09
	Nc-5-50-4	3,329.45	60.47	502.13	40.05	6,765.20	404.11	405.15	0.26
	Nc-5-50-5	3,569.70	89.64	797.54	41.35	8,527.40	418.34	419.73	0.33
	Nc-5-70-1	LIMIT	194.52	2,002.26	52.95	7,059.35	622.95	623.70	0.12
	Nc-5-70-2	LIMIT	188.64	2,014.85	56.15	7,446.90	604.73	609.01	0.71
	Nc-5-70-3	LIMIT	197.49	1,972.58	53.75	7,232.15	480.52	481.23	0.15
	Nc-5-70-4	LIMIT	177.31	1,674.19	50.95	6,884.50	478.36	480.45	0.44
	Nc-5-70-5	LIMIT	170.55	1,625.74	51.45	6,832.15	479.64	480.29	0.14

high percentage of task-insurance instances do not necessarily appear in the optimal STIP solution. For instance, in Nc-3-30-1, arc 257 is insured in 76 scenarios (more than any other arc) but is not insured in the optimal STIP solution; in fact, none of the four most frequently insured arcs in the row for Nc-3-30-1 are insured in the optimal STIP solution. However, Nc-5-30-1 displays a stronger correlation, in which three of the top four most frequently insured arcs are insured in the optimal STIP solution.

4.1.4. Analysis of the Critical Path Length Distribution. We also analyze the distribution of critical path lengths given different forms of the penalty function. In this experiment, we obtain an optimal solution x^*, compute $\psi^s(x^*)$ for each scenario $s \in \Omega$, and approximate the distribution of the critical path lengths with respect to different penalty functions. We consider the first 50-node graph in our data set, use a sample size of $N = 200$, and examine various two-segment continuous piecewise-linear penalty functions. Each penalty function has slope $m_1 = 1$ for the first segment, which has a penalty of zero when the critical path length is zero. The second piece of the function begins when the critical path length equals 2,300 (with a penalty of 2,300), and has slope m_2. We consider ratio values m_1/m_2 in the

Figure 3 Arc-Insuring Trend with Increasing Values of $c/(\bar{d}-\bar{g})$



	10.00%	20.00%	30.00%	40.00%	50.00%	60.00%	70.00%	80.00%	90.00%	100.00%
C-5-30	15.79%	10.53%	15.79%	10.53%	10.53%	10.53%	15.79%	10.53%	0.00%	0.00%
C-5-50	12.50%	18.75%	12.50%	6.25%	31.25%	6.25%	12.50%	0.00%	0.00%	0.00%
Nc-5-30	21.43%	7.14%	7.14%	7.14%	14.29%	28.57%	7.14%	7.14%	0.00%	0.00%
Nc-5-50	14.29%	14.29%	19.05%	23.81%	19.05%	4.76%	0.00%	4.76%	0.00%	0.00%

set $\{0.5, 0.75, 1, 1.25, 1.5\}$, so that the first two ratios give a nonlinear convex penalty function, the ratio of 1 yields a linear penalty function, and the last two values give a nonlinear concave penalty function. We optimize STIP given each penalty function, and plot distributions of the resulting critical path lengths over the 200 scenarios in Figure 4. In particular, for each horizontal segment labeled with value t, the plot gives the percentage of 200 scenarios that have critical path length in the interval $[t, t+50)$.

Figure 4 demonstrates that for functions having smaller ratios of m_1/m_2, the critical path lengths tend to be shorter on average, and exhibit smaller standard deviations compared to the critical path length distributions for functions having larger ratios of m_1/m_2. The actual means for critical path lengths given penalty functions having m_1/m_2 ratios of 0.5, 0.75, 1, 1.25, and 1.5 are 2,247, 2,328, 2,319, 2,409, and 2,483, respectively, and their standard deviations are 90, 132, 152, 158, and 215, respectively. This result is intuitive, because a sharply increasing penalty function essentially acts as a barrier function and limits the number of scenarios in which the critical path length is allowed to become large. By contrast, when m_1/m_2 is large, and the marginal cost of finishing the project very late becomes relatively small, the distribution functions tend to spread out across the spectrum of possible completion times.

Table 6 Illustration of First-Stage Solution Persistency

Arc no.	5	129	169	197	209	257	321	329	331
Nc-3-30-1	50	52	56	10	54[a]	76	58	67	47[a]
Nc-5-30-1	53	81	121[a]	5	153[a]	97	33	93[a]	65

[a]Arc was insured at optimality.

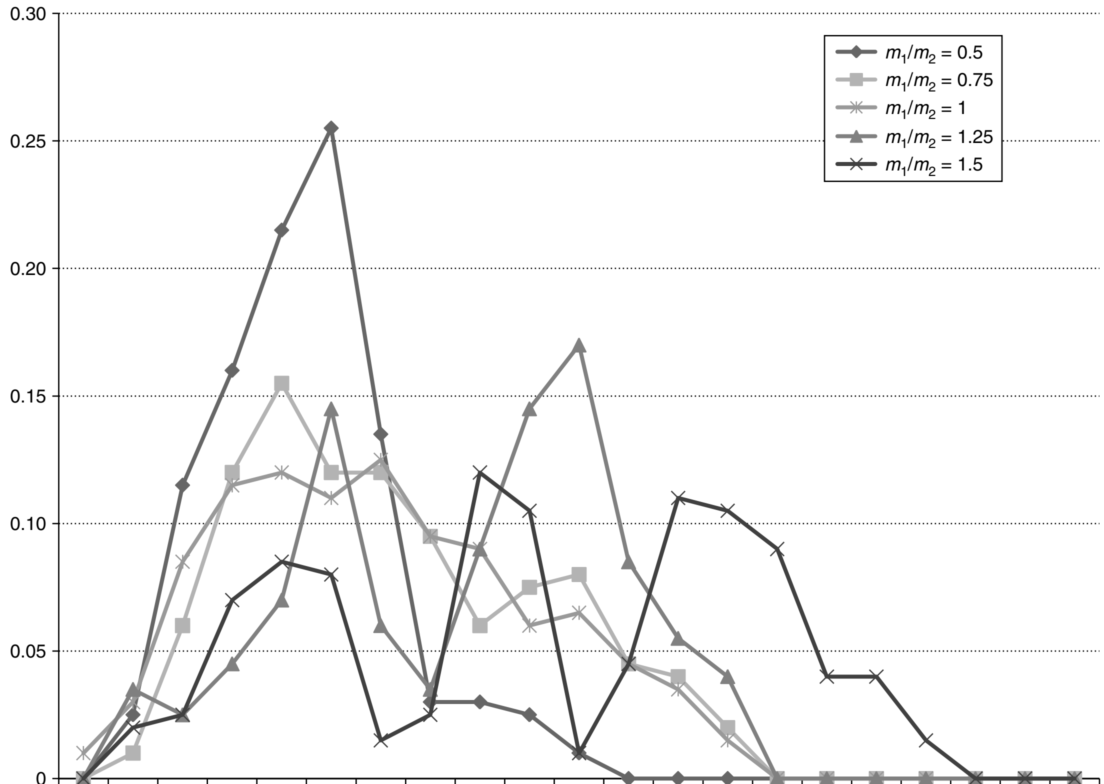
4.2. Chance-Constrained Formulation Case

Regarding problem CC_ϵ, when $\epsilon = 0$, one can use a scenario approximation method to solve $CC_{\epsilon=0}$ by solving the following approximation problem based on an independent Monte Carlo sample of random vectors $\xi^1, \ldots, \xi^N$:

$$\min\left\{\sum_{(i,j)\in\mathcal{A}} c_{ij}x_{ij}\colon x\in X, \psi(x,\xi^s)\le \mathcal{T}^s \quad \forall\, s=1,\ldots,N\right\}. \tag{38}$$

Luedtke and Ahmed (2008) approximate CC_ϵ with a general $\epsilon \ge 0$ by solving a sample approximation

Figure 4 Distributions of the Critical Path Lengths Given Different Penalty Functions



	2,000	2,050	2,100	2,150	2,200	2,250	2,300	2,350	2,400	2,450	2,500	2,550	2,600	2,650	2,700	2,750	2,800	2,850	2,900	2,950	3,000
$m_1/m_2 = 0.5$	0	0.025	0.115	0.16	0.215	0.255	0.135	0.03	0.03	0.025	0.01	0	0	0	0	0	0	0	0	0	0
$m_1/m_2 = 0.75$	0	0.01	0.06	0.12	0.155	0.12	0.12	0.095	0.06	0.075	0.08	0.045	0.04	0.02	0	0	0	0	0	0	0
$m_1/m_2 = 1$	0.01	0.03	0.085	0.115	0.12	0.11	0.125	0.095	0.09	0.06	0.065	0.045	0.035	0.015	0	0	0	0	0	0	0
$m_1/m_2 = 1.25$	0	0.035	0.025	0.045	0.07	0.145	0.06	0.035	0.09	0.145	0.17	0.085	0.055	0.04	0	0	0	0	0	0	0
$m_1/m_2 = 1.5$	0	0.02	0.025	0.07	0.085	0.08	0.015	0.025	0.12	0.105	0.01	0.045	0.11	0.105	0.09	0.04	0.04	0.015	0	0	0

problem. Let $\xi^1, \ldots, \xi^N$ be an independent Monte Carlo sample of the random vector ξ, and for a fixed $\alpha \in [0, 1)$, consider the following sample approximation problem:

$$\mathbf{CC}_\alpha^N: \quad \hat{z}_\alpha^N = \min\left\{ \sum_{(i,j)\in\mathcal{A}} c_{ij} x_{ij}\colon x \in X_\alpha^N \right\}, \tag{39}$$

where

$$X_\alpha^N = \left\{ x \in X\colon \frac{1}{N}\sum_{s=1}^{N} \mathbb{I}(\psi(x, \xi^s) \le \mathcal{T}^s) \ge 1 - \alpha \right\}. \tag{40}$$

Given ξ^s and first-stage binary variables $\hat{x}$, $\psi(\hat{x}, \xi^s)$ is given by $\mathrm{CPM}^s(\hat{x})$. Note that when $\alpha = 0$, the sample approximation problems (39) and (40) are equivalent to the scenario approximation program (38). We examine in §4.2.1 the case in which CC_α^N yields feasible solutions for CC_ϵ, and then discuss in §4.2.2 how to determine lower bounds with different confidences when $\alpha = \epsilon$.

Here we only test the first instance of each graph size, named as $\mathbf{y-1}$, where $\mathbf{y}$ represents the number of nodes (30, 50, or 70). We generate $\mathcal{T}^s$ from a uniform integer distribution over the interval $[0.7\hat{u}_n^s, 0.9\hat{u}_n^s]$. We again generate integer arc-insurance costs c_{ij}, $(i, j) \in \mathcal{A}$, uniformly over the interval [25, 50].

4.2.1. Feasible Solutions for $\mathbf{CC}_\alpha^N$. For a fixed value of $\alpha < \epsilon$, we wish to obtain a feasible solution to CC_ϵ with probability at least $1 - \delta$, for $\delta \in (0, 1)$. Because our feasible region X is finite, the result of Theorem 5 in Luedtke and Ahmed (2008) shows that it is sufficient to find a feasible solution to CC_α^N satisfying

$$N \ge \frac{1}{2(\alpha - \epsilon)^2} \log\frac{1}{\delta} + \frac{n}{2(\alpha - \epsilon)^2} \log U, \tag{41}$$

where in particular, when $\alpha = 0$, Theorem 7 suggests a sample size of

$$N \ge \frac{1}{\epsilon} \log\frac{1}{\delta} + \frac{n}{\epsilon} \log U, \tag{42}$$

and U is such that the number of feasible solutions obeys $|X| \le U^n$. In this case, because $x \in X \subseteq \{0, 1\}^n$, we use $U = 2$.

Table 7 Feasible Solution Results for CC_ϵ Sample Problems with $\epsilon = 0.01$

				Solution violation			Feasible solution costs			
α	$N = \|\Omega\|$	Instances	t_{avg}	Max	Min	Avg.	Num	Max	Min	Avg.
$\alpha = 0$	50	30-1	0.27	0.103	0.003	0.037	2	159	154	156.5
		50-1	0.42	0.124	0.007	0.041	1	212	212	212
		70-1	1.23	0.152	0.012	0.052	0	—	—	—
	80	30-1	0.39	0.052	0.000	0.019	3	159	154	156.33
		50-1	0.67	0.076	0.000	0.024	2	212	209	210.50
		70-1	1.78	0.092	0.001	0.031	2	234	232	233.00
	100	30-1	0.51	0.032	0.000	0.012	4	153	150	151.00
		50-1	0.78	0.038	0.000	0.017	4	209	207	208.00
		70-1	1.94	0.041	0.000	0.022	3	234	230	232.00
$\alpha = \epsilon$	1,000	30-1	2.58	0.041	0.000	0.012	6	143	140	141.17
		50-1	8.98	0.057	0.002	0.015	4	195	190	192.25
		70-1	63.99	0.023	0.005	0.013	4	223	220	221.50
	2,000	30-1	8.29	0.018	0.000	0.004	8	140	140	140.00
		50-1	23.81	0.041	0.001	0.008	7	195	190	190.71
		70-1	93.28	0.023	0.000	0.007	7	223	220	221.29

—: Not applicable.

Table 8 Feasible Solution Results for CC_ϵ Sample Problems with $\epsilon = 0.005$

				Solution violation			Feasible solution costs			
α	$N = \|\Omega\|$	Instances	t_{avg}	Max	Min	Avg.	Num	Max	Min	Avg.
$\alpha = 0$	100	30-1	0.51	0.032	0.000	0.012	2	153	151	152.00
		50-1	0.78	0.038	0.000	0.017	2	209	209	209.00
		70-1	1.94	0.041	0.000	0.022	1	234	234	234.00
	150	30-1	0.68	0.030	0.000	0.010	2	153	151	152.00
		50-1	1.32	0.031	0.000	0.009	3	209	207	208.33
		70-1	3.29	0.037	0.000	0.014	2	234	234	234.00
	200	30-1	1.21	0.025	0.000	0.007	4	153	151	151.50
		50-1	2.06	0.029	0.000	0.007	6	210	209	209.17
		70-1	5.46	0.032	0.000	0.009	5	234	234	234.00
$\alpha = \epsilon$	1,000	30-1	4.13	0.013	0.000	0.006	7	151	151	151.00
		50-1	9.36	0.021	0.000	0.013	9	205	203	203.22
		70-1	109.23	0.017	0.001	0.004	8	227	227	227.00
	2,000	30-1	11.09	0.007	0.000	0.002	9	151	151	151.00
		50-1	27.28	0.004	0.000	0.001	10	203	203	203.00
		70-1	132.56	0.008	0.001	0.002	9	227	227	227.00

We choose $\delta = 0.001$, $M = 10$ samples, and a reference sample size of $N' = 10{,}000$ scenarios for all three instances. We consider the cases of $\epsilon = 0.01$ and $\epsilon = 0.005$, and use $\alpha = 0$ and $\alpha = \epsilon$. Based on (41) and (42), when $\alpha = 0$, we use $N = 50$, 80, and 100 for the case with $\epsilon = 0.01$, and $N = 100$, 150, and 200 for the case with $\epsilon = 0.005$. When $\alpha = \epsilon$, we use sample sizes of $N = 1{,}000$ and $2{,}000$ scenarios. Next, we define the *violation risk* of a solution $\hat{x}$ as the percentage of reference scenarios for which the critical path (given insurance decisions $\hat{x}$) exceeds $\mathcal{T}^s$. We say that a solution is feasible if its violation risk does not exceed ϵ. Tables 7 and 8 report the objective function values of generated feasible solutions. In these tables t_{avg} and "Num" denote the average solution time and the total number of feasible solutions over all 10 samples of each instance, respectively.

In Table 7, considering instance 30-1, when $\alpha = 0$ and $N = 50$, the minimum violation among all solutions given by the $M = 10$ samples is $0.003 < \epsilon = 0.01$, and thus it is a feasible solution. The maximum violation risk among these samples is 0.103. Two out of 10 samples yield feasible solutions with objective values 159 and 154. With N increasing to 80 and 100, the number of feasible solutions increases to three and four, respectively. When $\alpha = \epsilon = 0.01$, $N = 1{,}000$, the minimum violation risk of all solutions is zero, and the maximum violation risk is 0.041 (which is not feasible). The number of feasible solutions increases to six. By setting $N = 2{,}000$, there are eight feasible solutions, all of which yield an objective value of 140. In Table 8, we set $\epsilon = 0.005$, and 30-1 yields more feasible solutions with better solution quality in each combination of α and N. Thus, by using $\alpha = \epsilon$, all instances yield more feasible solutions compared with the case of using $\alpha = 0$. However, more computational time is required to solve each instance's samples, because (41) requires larger values of N when $\alpha = \epsilon$. On the other hand, by decreasing the value of ϵ, we obtain more feasible solutions with higher solution quality in each setting.

4.2.2. Lower Bounds for CC_α^N. Theorem 4 of Luedtke and Ahmed (2008) provides a mechanism for obtaining a lower bound on CC_ϵ by solving CC_α^N with confidence $1 - \delta$. Given $\alpha \in [0, 1)$, we must choose positive integers N, L, and M such that $L \leq M$, and

$$\sum_{i=0}^{L-1} \binom{M}{i} \rho(\alpha, \epsilon, N)^i (1 - \rho(\alpha, \epsilon, N))^{M-i} \leq \delta, \qquad (43)$$

where $\rho(\alpha, \epsilon, N)$ represents the probability of having at most $\lfloor \alpha N \rfloor$ "successes" in N independent trials, in which the probability of a success in each trial is ϵ. With $\alpha = \epsilon$, one can choose the value of M independent of N to obtain a lower bound with confidence $1 - \delta$. Recalling that $M = 10$, if we take $L = 1$ (which corresponds to taking the minimum optimal solution over all $M = 10$ total runs not only over the feasible solutions), we obtain a lower bound with $1 - \delta = 0.999$ confidence. More generally, one can take a larger $L \in \{1, \ldots, M\}$ resulting in a lower bound with less confidence, but narrowing the optimality gap.

In Tables 9 and 10 we obtain lower bounds for the chance-constrained problems by taking $L = 1, 2, 3, 4, 5$, yielding corresponding confidence levels at least 0.999, 0.989, 0.945, 0.828, 0.623. We use the minimum objective function value of all feasible solutions to serve as the upper bound, and report the gaps between the upper and lower bounds with respect to each confidence.

Table 9 shows that for instance 30-1, setting $\alpha = \epsilon = 0.01$, $N = 1{,}000$, and $L = 1$, the minimum objective

Table 9 **Lower Bound Results for CC_ϵ Sample Problems with $\epsilon = 0.01$**

				Lower bound with confidence at least					Gap (%) with confidence at least				
α	$N = \|\Omega\|$	Instances	Upper bound	0.999	0.989	0.945	0.828	0.623	0.999	0.989	0.945	0.828	0.623
$\alpha = \epsilon$	1,000	30-1	140	137	137	139	139	140	2.19	2.19	0.72	0.72	0.00
		50-1	190	186	188	188	188	189	2.15	1.06	1.06	1.06	0.53
		70-1	220	217	217	219	219	219	1.38	1.38	0.46	0.46	0.46
	2,000	30-1	140	139	139	140	140	140	0.72	0.72	0.00	0.00	0.00
		50-1	190	188	189	189	190	190	1.06	0.53	0.53	0.00	0.00
		70-1	220	219	219	219	220	220	0.46	0.46	0.46	0.00	0.00

Table 10 **Lower Bound Results for CC_ϵ Sample Problems with $\epsilon = 0.005$**

				Lower bound with confidence at least					Gap (%) with confidence at least				
α	$N = \|\Omega\|$	Instances	Upper bound	0.999	0.989	0.945	0.828	0.623	0.999	0.989	0.945	0.828	0.623
$\alpha = \epsilon$	1,000	30-1	151	147	149	149	151	151	2.72	1.34	1.34	0.00	0.00
		50-1	203	201	203	203	203	203	1.00	0.00	0.00	0.00	0.00
		70-1	227	225	225	227	227	227	0.89	0.89	0.00	0.00	0.00
	2,000	30-1	151	149	151	151	151	151	1.34	0.00	0.00	0.00	0.00
		50-1	203	203	203	203	203	203	0.00	0.00	0.00	0.00	0.00
		70-1	227	225	225	225	225	225	0.89	0.89	0.89	0.89	0.89

function value, 137, over all 10 samples serves as a lower bound with confidence at least 0.999. We use the minimum objective function value 140 of all *feasible* solutions given in Table 7 as the upper bound, and the optimality gap is given as 2.19%. In Table 10, when $\alpha = \epsilon = 0.005$, we close the optimality gap by choosing $L = 4$ (with at least 0.828 confidence). This result is consistent with the fact that we have seven feasible solutions out of 10 samples in Table 8, and when $L > 10 - 7 = 3$, the lower and upper bounds are equal.

Similar to the discussion of feasible solutions, we narrow the optimality gap faster by using larger values of N. Furthermore, by allowing smaller ϵ, we can close the optimality gap with a higher confidence, which is intuitive because $\epsilon = 0.005$ yields more feasible solutions for each instance than $\epsilon = 0.01$. For example, when $\epsilon = 0.01$ and $N = 2{,}000$, Table 7 shows that we find eight feasible solutions out of 10 samples for instance 30-1. In Table 9, the two infeasible solutions both provide a lower bound of 139 to the original problem. We then set $L = 3$ and claim that 140 is a lower bound with confidence at least 0.945, which eliminates the optimality gap. In Table 10, we claim optimality with higher confidence: when $N = 2{,}000$ we close the optimality gap with confidence at least 0.989 for instance 30-1, compared with 0.945 in the case of $\epsilon = 0.01$ in Table 9.

5. Conclusion

In this paper, we consider a two-stage stochastic integer programming formulation for the STIP. The problem is naturally decomposable for convex penalty function problems. We employ RLT to make the STIP amenable to Benders decomposition given piecewise-linear nonconvex functions. Rather than explicitly computing the dual values of the resulting formulation, we propose an algorithm that quickly recovers all coefficients of Benders cuts based on the solution of a single critical path problem. We examine alternative dual optimal solutions to the dual problem, which yield alternative cuts, and we expand our decomposition technique to handle general lower semicontinuous penalty functions. We also cast the STIP in the context of a chance-constrained optimization problem and provide a cutting-plane algorithm for its solution. Future research may focus on comparing the computational efficiency of using different diversifications of Benders cut (18), and investigating sophisticated upper bounding algorithms for the STIP.

Acknowledgments

The authors are grateful to three anonymous referees, the associate editor, and the department editor, whose remarks led to an improved version of this paper. J. Cole Smith gratefully acknowledges the support of the Air Force Office of Scientific Research under Grants FA9550-07-1-0404 and FA9550-08-1-0189. Shabbir Ahmed gratefully acknowledges the support of the Air Force Office of Scientific Research under Grant FA9550-08-1-0117.

Appendix A. SP-LS$^s(\hat{x})$-RLT Simplification

We apply special structures RLT (Sherali et al. 1998) to SP-LS$^s(\hat{x})$, and decompose the formulation. The resulting subproblem is given by

$$\min \sum_{k=1}^{K^s} f_k^s \quad \text{(A1)}$$

$$\text{subject to: } -m_l^s v_{nk}^s + (Q - b_l^s) z_k^s \geq 0 \quad \forall\, l, k = 1, \ldots, K^s, l \neq k \quad \text{(A2)}$$

$$-m_k^s v_{nk}^s - b_k^s z_k^s + f_k^s \geq 0 \quad \forall\, k = 1, \ldots, K^s \quad \text{(A3)}$$

$$v_{nk}^s \geq 0 \quad \forall\, k = 1, \ldots, K^s \quad \text{(A4)}$$

$$v_{nk}^s - \tau_k^s z_k^s \geq 0 \quad \forall\, k = 1, \ldots, K^s \quad \text{(A5)}$$

$$-v_{nk}^s + (\tau_{l+1}^s + Q) z_k^s \geq 0 \quad \forall\, l, k = 1, \ldots, K^s, l \neq k \quad \text{(A6)}$$

$$-v_{nk}^s + \tau_{k+1}^s z_k^s \geq 0 \quad \forall\, k = 1, \ldots, K^s \quad \text{(A7)}$$

$$v_{jk}^s - v_{ik}^s - d_{ij}^s z_k^s + (d_{ij}^s - g_{ij}^s) w_{ijk}^s \geq 0 \quad \forall\, (i,j) \in \mathcal{A}, k = 1, \ldots, K^s \quad \text{(A8)}$$

$$-\sum_{k=1}^{K^s} w_{ijk}^s = -\hat{x}_{ij} \quad \forall\, (i,j) \in \mathcal{A} \quad \text{(A9)}$$

$$-w_{ijk}^s + z_k^s \geq 0 \quad \forall\, (i,j) \in \mathcal{A}, k = 1, \ldots, K^s \quad \text{(A10)}$$

$$w_{ijk}^s \geq 0 \quad \forall\, (i,j) \in \mathcal{A}, k = 1, \ldots, K^s \quad \text{(A11)}$$

$$\sum_{k=1}^{K^s} z_k^s = 1 \quad \text{(A12)}$$

$$f_k^s \geq 0 \quad \forall\, k = 1, \ldots, K^s \quad \text{(A13)}$$

$$z_k^s \geq 0 \quad \forall\, k = 1, \ldots, K^s. \quad \text{(A14)}$$

First, we observe that (A2) is implied by (A14) if $m_l^s = 0$, and otherwise ($m_l^s > 0$), (A2) is implied by (A7), because $(Q - b_l^s)/m_l^s > \tau_{k+1}^s$. Next, note that (A4), (A6), and (A13) are implied by (A5), (A7), and (A3), respectively. We thus obtain the revised SP-LS$^s(\hat{x})$-RLT as in §3.1.1.

Appendix B. Proof of Proposition 1

Proof. We demonstrate that the proposed dual solution is dual feasible and complementary slack to the primal solution. The dual feasibility conditions to SP-LS$^s(\hat{x})$-RLT are given by

$$-m_k^s A_k + B_k^+ - B_k^- + \sum_{i \in RS(n)} C_{ink} = 0 \quad \forall\, k = 1, \ldots, K^s, \quad \text{(B1)}$$

$$\sum_{l \in RS(i)} C_{lik} - \sum_{j \in FS(i)} C_{ijk} = 0 \quad \forall\, i = 1, \ldots, n-1, k = 1, \ldots, K^s, \quad \text{(B2)}$$

$$(d_{ij}^s - g_{ij}^s) C_{ijk} - D_{ij} - E_{ijk} \leq 0 \quad \forall\, (i,j) \in \mathcal{A}, k = 1, \ldots, K^s, \quad \text{(B3)}$$

$$A_k = 1 \quad \forall\, k = 1, \ldots, K^s, \quad \text{(B4)}$$

$$-b_k^s A_k - \tau_k^s B_k^+ + \tau_{k+1}^s B_k^- - \sum_{(i,j) \in \mathcal{A}} d_{ij}^s C_{ijk} + \sum_{(i,j) \in \mathcal{A}} E_{ijk} + F \leq 0 \quad \forall\, k = 1, \ldots, K^s, \quad \text{(B5)}$$

$$A_k, B_k^+, B_k^-, C_{ijk}, E_{ijk} \geq 0 \quad \forall\, (i,j) \in \mathcal{A}, k = 1, \ldots, K^s, \quad \text{(B6)}$$

where (B1), (B2), (B3), (B4), and (B5) are associated with primal v_{nk}, v_{ik} ($\forall\, i = 1, \ldots, n-1$), w_{ijk}, f_k, and z_k variables, respectively. We first verify the feasibility of our given dual values. Note that (B4) is directly satisfied by our choice of $A_k = 1$, $\forall\, k$. Also, for each k, all nonzero C_{ijk} values are equal and correspond to arcs (i,j) in a critical path, which verifies feasibility with respect to (B2). Furthermore, since $\sum_{i \in RS(n)} C_{ink} = C_{i^*nk}$, where $(i^*, n) \in \hat{\mathcal{A}}$, and because we explicitly set B_k^+ and B_k^- such that

$$C_{i^*nk} = m_k^s - B_k^+ + B_k^- \quad \forall\, k = 1, \ldots, K^s,$$

then (B1) is satisfied. Also, all A_k, B_k^+, B_k^-, and C_{ijk} values are clearly nonnegative, because ρ_k^L, ρ_k^R, and m_k^s are all nonnegative.

We next show that (B3) is satisfied, and that all E_{ijk} values are nonnegative. First, for $(i,j) \in \widehat{X}^1$, E_{ijk} is set to $(d_{ij}^s - g_{ij}^s) \cdot C_{ijk} - D_{ij}$ so that (B3) holds as an equality. To show that $E_{ijk} \geq 0$ in this case, for $k = 1, \ldots, k'$, we have $C_{ijk} = \max\{\rho_k^L, \rho_k^R, m_{k'}^s\} \geq m_{k'}^s$, and because $D_{ij} = \min_{k=k', \ldots, K^s}\{(d_{ij}^s - g_{ij}^s) C_{ijk}\} \leq (d_{ij}^s - g_{ij}^s) C_{ijk'}$ for $(i,j) \in \widehat{X}^1$, we get $(d_{ij}^s - g_{ij}^s) C_{ijk} \geq (d_{ij}^s - g_{ij}^s) \cdot m_{k'}^s \geq D_{ij}$, which verifies that $E_{ijk} = (d_{ij}^s - g_{ij}^s) C_{ijk} - D_{ij} \geq 0$. For $k = k'+1, \ldots, K^s$, $(d_{ij}^s - g_{ij}^s) C_{ijk} \geq \min_{k=k', \ldots, K^s}\{(d_{ij}^s - g_{ij}^s) C_{ijk}\} = D_{ij}$, which also yields $E_{ijk} = (d_{ij}^s - g_{ij}^s) C_{ijk} - D_{ij} \geq 0$.

Now, for $(i,j) \in \widehat{X}^0$, if $k = 1, \ldots, k'$, $(d_{ij}^s - g_{ij}^s) C_{ijk} \leq \max_{k=1,\ldots,k'}\{(d_{ij}^s - g_{ij}^s) C_{ijk}\} = D_{ij}$, whereas if $k = k'+1, \ldots, K^s$, $(d_{ij}^s - g_{ij}^s) C_{ijk} = (d_{ij}^s - g_{ij}^s) \min\{\rho_k^L, \rho_k^R, m_{k'}^s\} \leq (d_{ij}^s - g_{ij}^s) m_{k'}^s \leq \max_{k=1,\ldots,k'}\{(d_{ij}^s - g_{ij}^s) C_{ijk}\} = D_{ij}$. Therefore, setting $E_{ijk} = 0$, $\forall\, (i,j) \in \widehat{X}^0$, $k = 1, \ldots, K^s$ satisfies the remaining inequalities of (B3), and ensures nonnegativity of E_{ijk}, $\forall\, (i,j) \in \mathcal{A}$, $k = 1, \ldots, K^s$.

Next, we verify that our dual solution is feasible to (B5). For all $k = 1, \ldots, K^s$, we have $E_{ijk} = 0 = (d_{ij}^s - g_{ij}^s) C_{ijk} \hat{x}_{ij}$, $\forall\, (i,j) \in \widehat{X}^0$, and $E_{ijk} = (d_{ij}^s - g_{ij}^s) C_{ijk} - D_{ij}$, $\forall\, (i,j) \in \widehat{X}^1$. Hence,

$$\sum_{(i,j) \in \mathcal{A}} E_{ijk} = \sum_{(i,j) \in \widehat{X}^0} E_{ijk} + \sum_{(i,j) \in \widehat{X}^1} E_{ijk} = \sum_{(i,j) \in \hat{\mathcal{A}}} (d_{ij}^s - g_{ij}^s) C_{ijk} \hat{x}_{ij} - \sum_{(i,j) \in \widehat{X}^1} D_{ij} \quad \forall\, k = 1, \ldots, K^s. \quad \text{(B7)}$$

Substituting $\sum_{(i,j) \in \mathcal{A}} E_{ijk}$ with (B7), and noting (B1) and (B4), $C_{ijk} = 0$, $\forall\, (i,j) \notin \hat{\mathcal{A}}$, $C_{ijk} = C_{i^*nk}$, $\forall\, (i,j) \in \hat{\mathcal{A}}$, and $\sum_{(i,j) \in \hat{\mathcal{A}}} (d_{ij}^s - (d_{ij}^s - g_{ij}^s)\hat{x}_{ij}) = \hat{u}_n^s$, we write (B5) as

$$-b_k^s - \tau_k^s B_k^+ + \tau_{k+1}^s B_k^- - \hat{u}_n^s C_{i^*nk} - \sum_{(i,j) \in \widehat{X}^1} D_{ij} + F \leq 0 \quad \forall\, k = 1, \ldots, K^s, \quad \text{or}$$

$$-b_k^s - \tau_k^s B_k^+ + \tau_{k+1}^s B_k^- - \hat{u}_n^s C_{i^*nk} + \hat{f}^s \leq 0 \quad \forall\, k = 1, \ldots, K^s. \quad \text{(B8)}$$

If $C_{i^*nk} \geq m_k^s$, with respect to (B1), we substitute $B_k^+ = 0$, $B_k^- = C_{i^*nk} - m_k^s$, and (B8) becomes

$$(\hat{f}^s - (m_k^s \tau_{k+1}^s + b_k^s)) - (\hat{u}_n^s - \tau_{k+1}^s) C_{i^*nk} \leq 0. \quad \text{(B9)}$$

For $k = 1, \ldots, k'-1$, $(\hat{u}_n^s - \tau_{k+1}^s) > 0$, we have $C_{i^*nk} \geq (\hat{f}^s - (m_k^s \tau_{k+1}^s + b_k^s))/(\hat{u}_n^s - \tau_{k+1}^s) = \rho_k^R$; for $k = k', \ldots, K^s$, if $(\hat{u}_n^s - \tau_{k+1}^s) < 0$, (B9) requires $C_{i^*nk} \leq \rho_k^R$. If $k = k'$, $\hat{u}_n^s = \tau_{k'+1}^s$, and (B9) becomes $\hat{f}^s - (m_{k'}^s \tau_{k'+1}^s + b_{k'}^s) - 0 C_{i^*nk'} = 0 \leq 0$. All of the above cases are satisfied by the defined value of C_{i^*nk}.

Similarly, if $C_{i^*nk} < m_k^s$, substituting $B_k^- = 0$, $B_k^+ = m_k^s - C_{i^*nk}$ in (B8) gives us

$$(\hat{f}^s - (m_k^s \tau_k^s + b_k^s)) - (\hat{u}_n^s - \tau_k^s) C_{i^*nk} \le 0. \tag{B10}$$

We have $(\hat{u}_n^s - \tau_k^s) > 0$ for $k = 1, \ldots, k'$ and $(\hat{u}_n^s - \tau_k^s) < 0$ for $k = k'+1, \ldots, K^s$, due to which we require $C_{i^*nk} \ge (\hat{f}^s - (m_k^s \tau_k^s + b_k^s))/(\hat{u}_n^s - \tau_k^s) = \rho_k^L$ and $C_{i^*nk} \le \rho_k^L$, respectively. The given value of C_{i^*nk} satisfies both of the requirements. In particular, when $k = k'$, $\rho_{k'}^L = \rho_{k'}^R = m_{k'}^s$, and from the above analysis, we note that $C_{i^*nk'} \le \rho_{k'}^R$ and $C_{i^*nk'} \ge \rho_{k'}^L$. Therefore, $C_{i^*nk'} = m_{k'}^s$, and (B5) holds as an equality. Thus, all dual feasibility conditions are satisfied.

Finally, we demonstrate that the dual solution is complementary slack to the primal. Recall that in the primal solution, we have $v_{ik}^s = u_i^s z_k^s$, $\forall\, i, k$, $f_k^s = 0$, $\forall\, k \ne k'$ (i.e., if $z_k^s = 0$) and $f_{k'}^s = \hat{f}^s = m_{k'}^s u_n + b_{k'}^s$. Then (A3) is always binding $\forall\, k$, and (A5) and (A7) are binding for any $k \ne k'$. Recall that when $k = k'$, we set $C_{ijk'} = m_{k'}^s$ and $B_{k'}^+ = B_{k'}^- = 0$, which ensures complementary slackness with respect to (A5) and (A7). Because $w_{ijk}^s = \hat{x}_{ij} z_k^s$, primal constraints (A8) are potentially not binding only for $(i, j) \notin \hat{\mathcal{A}}$, and we satisfy complementary slackness by setting $C_{ijk} = 0$, $\forall\, k$, corresponding to $(i, j) \notin \hat{\mathcal{A}}$. Next, note that $E_{ijk} > 0$ only when $\hat{x}_{ij} = 1$ (i.e., $(i, j) \in \widehat{X}^1$) and thus $w_{ijk}^s = z_k^s$, ensuring that (A10) is binding. Now, with respect to dual inequality (B3), note that whenever its corresponding primal variable $w_{ijk}^s = 1$, we must have $k = k'(z_k^s = 1)$ and $\hat{x}_{ij} = 1$, in which case we set $E_{ijk'} = (d_{ij}^s - g_{ij}^s) C_{ijk'} - D_{ij}$, and ensure that (B3) is binding in this case. Finally, the primal variables z_k^s associated with dual inequality (B5) are all zero except for $z_{k'}^s = 1$, in which case we demonstrated in the proof of dual feasibility that (B5) is binding.

Hence, our dual solution is feasible and complementary slack to the primal feasible solution, which completes the proof. □

Appendix C. Proof of Proposition 2

Proof. We can verify the dual feasibility of our solution to constraints (B1), (B2), (B4), and the nonnegativity of A_k, B_k^+, B_k^-, C_{ijk} values, using the same arguments in the proof of Proposition 1.

We next demonstrate that (B3) is satisfied and all E_{ijk} are nonnegative. For arc $(i, j) \in \widehat{X}^1$, if $k = 1, \ldots, k'$, E_{ijk} is set to $(d_{ij}^s - g_{ij}^s) C_{ijk} - D_{ij}$, and (B3) is satisfied as an equality. Also, we have $(d_{ij}^s - g_{ij}^s) C_{ijk} = (d_{ij}^s - g_{ij}^s) \max\{\rho_k^L, \rho_k^R, m_{k'}^s\} \ge (d_{ij}^s - g_{ij}^s) m_{k'}^s \ge D_{ij}$, and $E_{ijk} \ge 0$ in this case using the same proof in Proposition 1. For $k = k'+1, \ldots, K^s$, because $C_{ijk} = E_{ijk} = 0$, we need to show that $D_{ij} \ge 0$. If $\sum_{(i,j) \in \widehat{X}^1} (d_{ij}^s - g_{ij}^s) m_{k'}^s \le \Delta^s(\hat{x})$, we have $D_{ij} = (d_{ij}^s - g_{ij}^s) m_{k'}^s \ge 0$, $\forall\, (i, j) \in \widehat{X}^1$; else, the values of D_{ij} are nonnegative for $h = 1, \ldots, h'$ such that $\sum_{h=1}^{h'} D_{(ij)_h} = \Delta^s(\hat{x})$, and $D_{(ij)_h} = 0$ for the remaining values of h, which verifies that $D_{ij} \ge 0$, $\forall\, (i, j) \in \widehat{X}^1$. Now, for all arcs $(i, j) \in \widehat{X}^0$ and $k = 1, \ldots, K^s$, note that $D_{ij} = \max_{k=1,\ldots,k'} \{(d_{ij}^s - g_{ij}^s) C_{ijk}\} \ge (d_{ij}^s - g_{ij}^s) C_{ijk}$ (because $C_{ijk} = 0, \forall\, (i, j) \in \hat{\mathcal{A}}, k = k'+1, \ldots, K^s$), i.e., $(d_{ij}^s - g_{ij}^s) \cdot C_{ijk} - D_{ij} \le 0$, and thus setting $E_{ijk} = 0$ satisfies (B3).

Next, we verify the feasibility of (B5). First, for $k = 1, \ldots, k'$, $E_{ijk} = 0 = (d_{ij}^s - g_{ij}^s) C_{ijk} \hat{x}_{ij}$, $\forall\, (i, j) \in \widehat{X}^0$, and $E_{ijk} = (d_{ij}^s - g_{ij}^s) C_{ijk} \hat{x}_{ij} - D_{ij}$, $\forall\, (i, j) \in \widehat{X}^1$. Hence, we have satisfied both (B7) and (B8) for $k = 1, \ldots, k'$. By using the same argument in the proof of Proposition 1, one can verify the feasibility of (B5) in the case of $k = 1, \ldots, k'$.

For $k = k'+1, \ldots, K^s$, $C_{ijk} = E_{ijk} = 0$, $\forall\, (i, j) \in \mathcal{A}$, and based on (B1), $B_k^+ = m_k^s$, $B_k^- = 0$, we now transform (B5) into

$$-b_k^s - \tau_k^s m_k^s + \hat{f}^s + \sum_{(i,j) \in \widehat{X}^1} D_{ij} \le 0, \quad \forall\, k = k'+1, \ldots, K^s. \tag{C1}$$

Note that by sequentially setting $D_{(ij)_h} = \min\{(d_{ij}^s - g_{ij}^s) m_{k'}^s, \Delta^s(\hat{x}) - \sum_{l=1}^{h-1} D_{ij_l}\}$, for all $(i, j) \in \widehat{X}^1$, $h = 1, \ldots, H$, we have

$$\sum_{(i,j) \in \widehat{X}^1} D_{ij} = \min\left\{ \sum_{(i,j) \in \widehat{X}^1} (d_{ij}^s - g_{ij}^s) m_{k'}^s, \Delta^s(\hat{x}) \right\} \le \Delta^s(\hat{x})$$
$$= \tau_{k'+1}^s m_{k'+1}^s + b_{k'+1}^s - \hat{f}^s,$$

which satisfies (C1) (because $\tau_{k'+1}^s m_{k'+1}^s + b_{k'+1}^s - \hat{f}^s \le \tau_k^s m_k^s + b_k^s - \hat{f}^s$, $\forall\, k = k'+1, \ldots, K^s$), and thus (B5) is also satisfied in the case of $k = k'+1, \ldots, K^s$.

Complementary slackness of the dual solution to the primal is verified by the same argument as in the proof of Proposition 1. This completes the proof. □

Appendix D. Proof of Proposition 3

Proof. We demonstrate the dual feasibility and complementary slackness of the proposed dual solution. First, the dual feasibility conditions to CC_ϵ-SP-$^s(\hat{x})$-RLT are given by

$$\sum_{l \in RS(n)} (C_{\text{ln}}^- - C_{\text{ln}}^+) - B^+ + B^- = 0, \tag{D1}$$

$$\sum_{l \in RS(i)} (C_{li}^- - C_{li}^+) - \sum_{l \in FS(i)} (C_{ij}^- - C_{ij}^+) = 0 \quad \forall\, i = 1, \ldots, n-1, \tag{D2}$$

$$\sum_{l \in RS(n)} C_{\text{ln}}^+ - B^- = 0, \tag{D3}$$

$$\sum_{l \in RS(i)} C_{li}^+ - \sum_{l \in FS(i)} C_{ij}^+ = 0 \quad \forall\, i = 1, \ldots, n-1, \tag{D4}$$

$$(d_{ij}^s - g_{ij}^s)(C_{ij}^- - C_{ij}^+) - D_{ij} - E_{ij}^- + E_{ij}^+ \le 0 \quad \forall\, (i, j) \in \mathcal{A}, \tag{D5}$$

$$-\sum_{(i,j) \in \mathcal{A}} d_{ij}^s C_{ij}^- + \sum_{(i,j) \in \mathcal{A}} d_{ij}^s C_{ij}^+ + \sum_{(i,j) \in \mathcal{A}} (E_{ij}^- - E_{ij}^+) + (\mathcal{T}^s + Q) B^+ - \mathcal{T}^s B^- \le 1, \tag{D6}$$

$$C_{ij}^-, C_{ij}^+, B^+, B^-, D_{ij}, E_{ij}^-, E_{ij}^+ \ge 0, \quad \forall\, (i, j) \in \mathcal{A}, \tag{D7}$$

where (D1), (D2), (D3), (D4), (D5), and (D6) are associated with primal v_n^s, v_i^s ($\forall\, i = 1, \ldots, n-1$), u_n^s, u_i^s ($\forall\, i = 1, \ldots, n-1$), w_{ij}^s, and p^s variables, respectively. First, note that all C_{ij}^- and C_{ij}^+ values are equal corresponding to $(i, j) \in \hat{A}$, and $C_{ij}^- = C_{ij}^+ = 0$ for $(i, j) \in \mathcal{A} \setminus \hat{\mathcal{A}}$, so that (D2) and (D4) are both satisfied. Constraint (D3) is satisfied by setting $B^- = C_{i^*n}^+$, where arc (i^*, n) is the critical path arc entering node n. We also satisfy (D1) by setting $B^+ = B^-$.

Constraint (D5), for $(i, j) \in \mathcal{A} \setminus \hat{\mathcal{A}}$, is satisfied by setting all C_{ij}^-, C_{ij}^+, D_{ij}, E_{ij}^-, and E_{ij}^+ values equal to zero. Recalling that $C_{ij}^- = C_{ij}^+$, $\forall\, (i, j) \in \hat{\mathcal{A}}$, we satisfy (D5) as an equality for $(i, j) \in \widehat{X}^1$ by setting $E_{ij}^+ = E_{ij}^- = (d_{ij}^s - g_{ij}^s)/(\hat{u}_n^s - \mathcal{T}^s)$ and $D_{ij} = 0$. For $(i, j) \in \widehat{X}^0$, we have (D5) equaling zero because

$E^+_{ij} = E^-_{ij} = D_{ij} = 0$. Also note that E^+_{ij}, E^-_{ij}, and D_{ij} are all nonnegative for all $(i, j) \in \mathcal{A}$.

We verify that our dual solution is feasible for (D6). Given dual values of E^+_{ij}, E^-_{ij}, C^-_{ij}, C^+_{ij}, B^+, and B^-, (D6) becomes $-\sum_{(i,j)\in\mathcal{A}} d^s_{ij}(C^-_{ij} - C^+_{ij}) + 0 + (\mathcal{T}^s + Q - \mathcal{T}^s)/(\hat{u}^s_n - \mathcal{T}^s) = 1$ by assigning the compensating value $\hat{u}^s_n - \mathcal{T}^s \geq 0$ to Q, which is satisfied as an equality.

Finally, we demonstrate that the duals are also complementary slack to the the primal. Because $\hat{p}^s = 1$, (29) becomes $u^s_j - u^s_i \geq d^s_{ij} - (d^s_{ij} - g^s_{ij})\hat{x}_{ij}$, $\forall (i, j) \in \mathcal{A}$, which are potentially not binding for $(i, j) \notin \hat{\mathcal{A}}$, and we satisfy the complementary slackness by setting $C^-_{ij} = 0$, corresponding to arcs $(i, j) \notin \hat{\mathcal{A}}$. Constraints (30) are always binding, and thus we require $C^+_{ij} \geq 0$, $\forall (i, j) \in \mathcal{A}$. For $D_{ij} = 0$ and $E^+_{ij} \geq 0, \forall (i, j) \in \mathcal{A}$, constraints (33) and (35) are always binding because whenever a Benders cut needed to be generated, we have $\hat{p}^s = 1$ and $\hat{w}^s_{ij} = \hat{x}_{ij}$. If $\hat{x}_{ij} = 0$, (i.e., $(i, j) \in \widehat{X}^0$), (34) holds as an inequality, and we have $E^-_{ij} = 0$ in this case. Next, (D5) is always binding for the given dual values, and for $\hat{p}^s = 1$, we have already shown that (D6) holds as an equality in the foregoing analysis. Therefore, the proposed values of duals are dual feasible and complementary slack to the primal. This completes the proof. □

Appendix E. Expectation-Based Sample Average Approximation

The SAA is an approach for solving stochastic optimization problems by using Monte Carlo simulation. A set of N sample scenarios $\omega^1, \ldots, \omega^N$ is generated from Ω according to its probability distribution. We then solve a deterministic optimization problem specified by scenarios $\omega^1, \ldots, \omega^N$. We approximate the expected second-stage recourse costs by the sample average function $(1/N)\sum_{n=1}^N F(x, \xi(\omega^n))$, and the overall optimal objective value as

$$f_N = \min_{x\in X} \left\{ \sum_{(i,j)\in\mathcal{A}} c_{ij}x_{ij} + \frac{1}{N}\sum_{n=1}^{N} F(x, \xi(\omega^n)) \right\}. \tag{E1}$$

Now suppose that we generate M independent sets of samples, each of size N, and solve problems (E1) independently for each set of samples. We denote f^m_N and $\hat{x}^m_N$ as the optimal objective value and solution to (E1), respectively, for each set of samples, $m = 1, \ldots, M$.

The average of the M optimal objective values, $\bar{f}_{N,M} = (1/M)\sum_{m=1}^M f^m_N$ provides a statistical estimate for the lower bound on the optimal objective function value.

Next, we pick any feasible first-stage solution from among optimal solutions $\hat{x} = \hat{x}^m_N$ to (E1), for some $m \in \{1, \ldots, M\}$. By fixing $x = \hat{x}$ in the second stage, we estimate the optimal objective value using a reference sample of size N'. We then compute an upper bound on the optimal objective function value as

$$\hat{f}_{N'}(\hat{x}) = \sum_{(i,j)\in\mathcal{A}} c_{ij}\hat{x}_{ij} + \frac{1}{N'}\sum_{n=1}^{N'} F(\hat{x}, \xi(\omega^n)). \tag{E2}$$

Because computing $\hat{f}_{N'}(\hat{x})$ requires only the solution of N' subproblems, we can generate the reference sample with size N' much larger than N, independent of the samples used in problems (E1). To obtain the best upper bound, we thus select

$$x^* \in \operatorname*{arg\,min}_{\hat{x}\in\{\hat{x}^1_N, \ldots, \hat{x}^M_N\}} \{\hat{f}_{N'}(\hat{x})\}, \tag{E3}$$

which yields an absolute optimality gap $\hat{f}_{N'}(x^*) - \bar{f}_{N,M}$.

References

Benders, J. F. 1962. Partitioning procedures for solving mixed variables programming problems. *Numerische Mathematik* **4**(1) 238–252.

Bertsimas, D., K. Natarajan, C.-P. Teo. 2006. Persistence in discrete optimization under data uncertainty. *Math. Programming, Ser. B* **108**(2–3) 251–274.

Bowman, R. A. 1995. Efficient estimation of arc criticalities in stochastic activity networks. *Management Sci.* **41**(1) 58–67.

Bowman, R. A., J. A. Muckstadt. 1993. Stochastic analysis of cyclic schedules. *Oper. Res.* **41**(5) 947–958.

Brucker, P., A. Drexl, R. Moehring, K. Neumann, E. Pesch. 1999. Resource-constrained project scheduling: Notation, classification, models, and methods. *Eur. J. Oper. Res.* **112**(1) 3–41.

Burt, J. M., M. B. Garman. 1971. Conditional Monte Carlo: A simulation technique for stochastic network analysis. *Management Sci.* **18**(3) 207–217.

Chen, X., M. Sim, P. Sun, J. Zhang. 2008. A linear-decision based approximation approach to stochastic programming. *Oper. Res.* **56**(2) 344–357.

Chtourou, H., M. Haouari. 2008. A two-stage-priority-rule-based algorithm for robust resource-constrained project scheduling. *Comput. Indust. Engrg.* **55**(1) 183–194.

Cormican, K. J., D. P. Morton, R. K. Wood. 1998. Stochastic network interdiction. *Oper. Res.* **46**(2) 184–197.

Demeulemeester, E. L., W. S. Herroelen. 2002. *Project Scheduling: A Research Handbook*. Springer, Norwell, MA.

Elmaghraby, S. E., A. A. Ferreira, L. V. Tavares. 2000. Optimal start times under stochastic activity durations. *Internat. J. Production Econom.* **64**(1–3) 153–164.

Goldratt, E. M. 1997. *Critical Chain*. North River Press, Great Barrington, MA.

Golenko-Ginzburg, D., A. Gonik. 1998. A heuristic for network project scheduling with random activity durations depending on the resource allocation. *Internat. J. Production Econom.* **55**(2) 149–162.

Golenko-Ginzburg, D., A. Gonik, S. Sitniakovski. 2000. Resource supportability model for stochastic network projects under a chance constraint. *Comm. Dependability Quality Management* **3**(1) 89–102.

Gutjahr, W. J., C. Strauss, E. Wagner. 2000. A stochastic branch-and-bound approach to activity crashing in project management. *INFORMS J. Comput.* **12**(2) 125–135.

Hagstrom, J. N. 1990. Computing the probability distribution of project duration in a PERT network. *Networks* **20**(2) 231–244.

Herroelen, W., R. Leus. 2001. On the merits and pitfalls of critical chain scheduling. *J. Oper. Management* **19**(5) 559–577.

Herroelen, W., R. Leus. 2005. Project scheduling under uncertainty: Survey and research potentials. *Eur. J. Oper. Res.* **165**(2) 289–306.

Hindelang, T. J., J. F. Muth. 1979. A dynamic programming algorithm for decision CPM networks. *Oper. Res.* **27**(2) 225–241.

Iida, T. 2000. Computing bounds on project duration distributions for stochastic PERT networks. *Naval Res. Logist.* **47**(7) 559–580.

ILOG. 2008. CPLEX 11.0 User's Manual. ILOG, Armonk, NY.

Janjarassuk, U., Linderoth, J. T. 2008. Reformulation and sampling to solve a stochastic network interdiction problem. *Networks* **52**(3) 120–132.

Kelley, J. E., Jr. 1961. Critical-path planning and scheduling: Mathematical basis. *Oper. Res.* **9**(3) 296–320.

Kelley, J. E., Jr. 1963. The critical path method: Resource planning and scheduling. *Indust. Scheduling* **2**(1) 347–365.

Kleywegt, A. J., A. Shapiro, T. Homem-de-Mello. 2001. The sample average approximation method for stochastic discrete optimization. *SIAM J. Optim.* **12**(2) 479–502.

Kulkarni, V. G., V. G. Adlakha. 1986. Markov and Markov-regenerative PERT networks. *Oper. Res.* **34**(5) 769–781.

Laslo, Z. 2003. Activity time-cost tradeoffs under time and cost chance constraints. *Comput. Indust. Engrg.* **44**(3) 365–384.

Luedtke, J., S. Ahmed. 2008. A sample approximation approach for optimization with probabilistic constraints. *SIAM J. Optim.* **19**(2) 674–699.

Mak, W. K., D. P. Morton, R. K. Wood. 1999. Monte Carlo bounding techniques for determining solution quality in stochastic programs. *Oper. Res. Lett.* **24**(1) 47–56.

Mitchell, G., T. Klastorin. 2007. An effective methodology for the stochastic project compression problem. *IIE Trans.* **39**(10) 957–969.

Moehring, R. H. 1984. Minimizing costs of resource requirements in project networks subject to a fixed completion time. *Oper. Res.* **32**(1) 89–120.

Norkin, V. I., G. Ch. Pflug, A. Ruszczyński. 1998. A branch-and-bound method for stochastic global optimization. *Math. Programming* **83**(1–3) 425–450.

Ozdamar, L., G. Ulusoy. 1995. A survey on the resource constrained project scheduling problem. *IIE Trans.* **27**(5) 574–586.

Patterson, J. H. 1984. A comparison of exact procedures for solving the multiple constrained resource, project scheduling problem. *Management Sci.* **30**(7) 854–867.

Scholl, A. 2001. *Robuste Planung und Optimierung: Grundlagen, Konzepte, und Methoden*. Physica-Verlag, Experimentelle Untersuchungen, Heidelberg, Germany.

Schultz, R. 2003. Stochastic programming with integer variables. *Math. Programming* **97**(1–2) 285–309.

Shapiro, A., T. Homem-de-Mello. 2000. On the rate of convergence of optimal solutions of Monte Carlo approximations of stochastic programs. *SIAM J. Optim.* **11**(1) 70–86.

Sherali, H. D. 2001. On mixed-integer zero-one representations for separable lower-semicontinuous piecewise-linear functions. *Oper. Res. Lett.* **28**(4) 155–160.

Sherali, H. D., W. P. Adams. 1990. A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. *SIAM J. Discrete Math.* **3**(3) 411–430.

Sherali, H. D., W. P. Adams. 1994. A hierarchy of relaxations and convex hull characterizations for mixed-integer zero-one programming problems. *Discrete Appl. Math.* **52**(1) 83–106.

Sherali, H. D., B. M. P. Fraticelli. 2002. A modification of Benders' decomposition algorithm for discrete subproblems: An approach for stochastic programs with integer recourse. *J. Global Optim.* **22**(1–4) 319–342.

Sherali, H. D., J. C. Smith. 2009. Two-stage stochastic risk threshold and hierarchical multiple risk problems: Models and algorithms. *Math. Programming, Ser. A* **120**(2) 403–427.

Sherali, H. D., W. P. Adams, P. J. Driscoll. 1998. Exploiting special structures in constructing a hierarchy of relaxations for 0-1 mixed integer problems. *Oper. Res.* **46**(3) 396–405.