A Solution Method for Single-Source Capacitated Facility Location Problems Under Decision-Dependent Demand Uncertainty

Giovanni Pantuso
Department of Mathematical Sciences
University of Copenhagen
Universitetsparken 5, 2100, Copenhagen, Denmark
gp@math.ku.dk
Miguel Lejeune
The George Washington University
Washington, United States
mlejeune@gwu.edu

Abstract

We present a model and an exact solution method for *single-source capacitated facility location problems* with decision-dependent demand uncertainty. The model allows the joint probability distribution of customer demands to depend on the facility to which demand points are allocated. The number of possible decision-dependent distributions is therefore equal to the number of feasible allocations of demand points to (groups of) facilities. To address the exponential size of the model, we apply and extend a recent version of the L-Shaped method for stochastic programs with decision-dependent uncertainty. In particular, we compare the original optimality cut with a new optimality cut based on a bilinear reformulation. We prove that, separated together with the original one, the new cut yields a tighter restriction, at the expense of introducing bilinear terms. In addition, we prove an effective decision-independent valid inequality that significantly improves convergence. Extensive numerical tests on 180 instances demonstrate that the method, supported by the valid inequality, can find optimal or near-optimal solutions to problem instances with up to 6^{20} potential probability distributions. The tests also show that, while the new cut closes slightly more instances, the original cut remains preferable in terms of solution time and optimality gap.

Keywords: Stochastic Programming, Endogenous Uncertainty, Facility Location, Decision-Dependent Uncertainty.

1 Introduction

Facility location decisions are significantly influenced by uncertainty [Liu et al., 2022, Correia and Saldanha-da Gama, 2019, Laporte et al., 2019, 1994, Louveaux, 1986, Snyder, 2006]. The

optimization literature is rich of methodology to address the resulting problems. The corresponding literature is thoroughly investigated and classified in the survey paper by Snyder [2006] and recent books [Correia and Saldanha-da Gama, 2019, Saldanha-da Gama and Wang, 2024].

In several real-life applications, location decision problems are further complicated by additional constraints. A prime example, which is investigated in this paper, is the *single-source* requirement. Upon making location decisions, the decision maker simultaneously assigns each customer to at most one capacitated facility and build the corresponding infrastructure. The allocated facility supplies its entire demand [Guastaroba and Speranza, 2014]. This requirement is rooted in several industrial applications. For instance, in healthcare planning, decision makers locate clinics and simultaneously assign population zones to a single clinic. Similarly, when locating retail stores, firms may assign geographical zones to a single retail, which covers its entire demand. In the telecommunications industry, service providers simultaneously decide where to place optical fiber aggregation nodes and which households to connect to each installed node. Each household is connected to exactly one node, requiring fixed investments in trenching, cabling, and installation. In water supply system planning, decision makers determine which treatment plants to build and assign each demand zone to a single plant. This assignment entails constructing pipelines and pressure-control infrastructure. The single-source constraint makes the problem significantly harder to solve, see Klose and Drexl [2005]. For example, assume that a set of open facilities has been chosen. If multiple sources are allowed, the decision problem of optimally supplying customers from those facilities forms a transportation problem. However, if single-sourcing is required, the corresponding assignment problem forms a special *Generalized Assignment Problem* (GAP) which is NP-hard, see Fisher et al. [1986].

In some applications, demand uncertainty may depend on the facility serving a given area. In healthcare planning, demand for healthcare services may vary with the accessibility and quality of the assigned clinic. Zones served by nearby or better-equipped clinics may exhibit higher utilization rates. In retail network design, the demand generated by a zone may depend on the attractiveness of the assigned store. Factors, such as distance, product assortment, and service quality, can all influence customer demand and therefore affect both location and assignment decisions. In the planning of emergency and public services, such as ambulance, vaccination, or social-service systems, utilization may increases when services become easier to access. Assigning a region to a closer facility may induce additional demand. Finally, in the planning of transportation systems, travelers may use a transit hub, station, or shared-mobility facility more often when assigned to a convenient or high-capacity location.

Despite the abundance of methodological approaches, the likely influence of location decisions on the uncertainties has not been much investigated. To our knowledge, the only available studies, which address the impact of location decisions on the uncertainty are those of Basciftci et al. [2021] and Pantuso [2025]. Basciftci et al. [2021] study a facility location problem where demand is endogenously influenced by location decisions. Particularly, the mean and variance of the demand distribution is a function of facility-location decisions. The authors propose a distributionally-robust optimization model to address misspecifications in the demand distri-

bution. That is, location decisions determine an ambiguity set, which contains the potential demand distributions that match the first two moments. The objective of the model is then to minimize the facility-location costs and the worst-case expected costs of transportation and unsatisfied demand. The authors show that the model can be reformulated as a mixed-integer linear programming (MILP) problem in the special case where demand mean and variance are piecewise linear functions of location decisions. In an extensive computational study, the authors demonstrate the superiority of the proposed model to traditional robust optimization and stochastic programs. Pantuso [2025] studies a facility location problem where all parameters of the problem may be affected by uncertainty. The author presents a general mathematical model in the form of a stochastic program with decision-dependent uncertainty. The central assumption is that the distribution of the data depends on the facilities (or groups of facilities) opened in the first-stage solution. A consequence of this is that there exist only finitely many possible distributions that may be induced, one for each feasible facility location decision. However, the number of feasible solutions – and corresponding distributions – is $\mathcal{O}(2^I)$ with I denoting the number of facilities. The author shows that the problem can be formulated as an MILP problem with exponentially many constraints. To solve the problem, the author proposed an exact decomposition method. Exact solutions are found for instances with up to 2^{17} distributions.

In the case of decision-dependent uncertainty, the distribution of demand can no longer be meaningfully assumed to depend on the choice of open facilities, as done in Pantuso [2025].

When each customer's demand is satisfied entirely by a single source, it is more realistic to assume that the demand distribution depends on the facility to which the customer is assigned. Consequently, if I denotes the number of facilities and J the number of customers, the number of allocation decisions – and corresponding distributions – is $\mathcal{O}(I+1)^J$, see Section 2. This number is, in general, much larger than 2^I as in the case of Pantuso [2025]. We address this case in this article.

At a more general level, stochastic programs with decision-dependent uncertainty are notoriously challenging. While the research literature offers several practical applications, see e.g., Viswanath et al. [2004], Peeta et al. [2010], Lejeune et al. [2024], Tong et al. [2012], Escudero et al. [2018], Lejeune and Ma [2025], Lejeune et al. [2026], general-purpose methodology remains sparse. In particular, Jonsbråten et al. [1998] study decision problems where decisions have an impact both on the probability of future random events, as well as on timing of their resolution. The authors assume that the set of probability measures which can be enforced by decisions is finite, as is the case in our paper. The authors show that, for moderately-sized instances, the problem can be solved by enumerating the possible distributions and solving the corresponding stochastic program. Pantuso [2021] extends the model of Jonsbråten et al. [1998] by allowing decisions at all stages to determine the probability measure. The author proposes a formulation that avoids explicit non-anticipativity constraints, see e.g., Apap and Grossmann [2017], Hooshmand and MirHassani [2016, 2018]. Hellemo et al. [2018] discuss several ways

to model the interplay between decisions and uncertainties. In their work, it is evident how the resulting formulations are typically nonlinear. Their characteristics heavily depend on the

nature of the relationship between decisions and distributions, severely complicating the development of general-purpose solution approaches. Finally, Pantuso and Hewitt [2025] propose an extension of the L-Shaped method for two-stage stochastic programs with decision-dependent uncertainty. The method is based on identifying the enforced distribution for each candidate first-stage solution prior to separate decision-dependent cuts. It relies on the assumption that there exists finitely many probability distributions that may be enforced. This assumption naturally holds for many combinatorial optimization problems, such as the one studied in this paper.

In this paper we consider and extend the *single-source capacitated stochastic facility location problem* presented by Laporte et al. [1994] (see also Hindi and Pieńkosz [1999], Ahuja et al. [2004], Lin [2009], Yang et al. [2012], Laporte et al. [2019] for variants of the problem). In this problem, each demand point may be assigned to at most one facility. Hence, first-stage decisions determine both the open facilities and the assignment of demand points to facilities. A fixed cost is paid upon opening a facility and creating the link between a demand point and a facility. Demand may be satisfied only through open links. Each point has a stochastic demand, which is revealed only after facilities and links have been open. For this reason, the demand does not need to be met exactly. Rather, a revenue per unit of satisfied demand is introduced. The setup of Laporte et al. [1994] is based on the assumption that the demand distribution is known and independent of decisions. We relax this (restrictive) assumption by allowing the probability distribution of the demands to depend on the facility to which the demand point is allocated. This, in turn, may allow modeling situations where demand is influenced by e.g., the distance to the provider, the cost born for each unit of demand satisfied, or characteristics of the providers. This generates a problem where there exists one probability distribution for each possible allocation of demand points to facilities. As we will see, this number is exponential in the number of customers.

The contribution of this paper is as follows:

- We propose a stochastic programming model for a *single-source capacitated facility location problem* under endogenous stochastic demand. The model allows the distribution of demand to depend on the allocation of demand points to facilities. The proposed model is general in the sense that it models a range of scenarios having the original problem presented by Laporte et al. [1994] at one extreme and, at the other extreme, a problem where there are as many distributions as there are allocations of demand points to facilities.

- To solve the problem, we present an exact cutting-plane algorithm, which can be regarded as a version of the L-Shaped method for stochastic programs with decisions-dependent uncertainty developed by Pantuso and Hewitt [2025]. In particular, we develop a new general optimality cut and prove that, separated together with the cut of Pantuso and Hewitt [2025], it yields a tighter restriction. The tightening requires the addition of bilinear terms, which can be linearized exactly. This, however, introduces a number of linearization variables that grows quadratically with the number of demand points. In addition, we prove a very tight, problem-specific, valid inequality, which significantly

increases the performance of the method.

- We present results of extensive numerical tests. These illustrate that the method can solve instances with up to 6^{20} potential probability distributions. They also show that the new cut closes slightly more instances than the original one, but the original cut remains preferable overall, both in terms of solution time and optimality gap attained. We explain that the bilinear terms make each round of separation more expensive leading to the generation of fewer cuts within the same time limit. This illustrates that a cut that is provably tighter is not necessarily more effective once the cost of separating it is taken into account.

In the remainder of this article we begin by presenting the model in Section 2. Following, we describe the solution method in Section 3. In Section 4, we present the results of numerical experiments. Finally, we draw conclusions in Section 5.

2 Model

Let $\mathcal{I}$ be the set of potential facility locations and $\mathcal{J}$ be the set of demand points to serve. Opening a facility at location i determines a fixed cost F_i. Opening a connection between facility i and demand point j determines a fixed cost C_{ij}. Each demand point has uncertain demand. We model the demand as a $|\mathcal{J}|$-dimensional real-valued random variable $\xi : (\Omega, \mathcal{F}) \to (\mathbb{R}^{|\mathcal{J}|}, \mathcal{B}^{|\mathcal{J}|})$ on a measurable space $(\Omega, \mathcal{F})$, where $\mathcal{B}^{|\mathcal{J}|}$ is the Borel σ-algebra on $\mathbb{R}^{|\mathcal{J}|}$. We allow the probability distribution of ξ on $\mathbb{R}^{|\mathcal{J}|}$ to depend on the location of the facility to which each demand point has been assigned.

Let $x := (x_i)_{i\in\mathcal{I}} \in \{0,1\}^{|\mathcal{I}|}$ denote the vector of facility location decisions and let $y := (y_{ij})_{i\in\mathcal{I},j\in\mathcal{J}} \in \{0,1\}^{|\mathcal{I}|\times|\mathcal{J}|}$ represent the assignments of demand points to facilities. To allow for a rather general setup, we assume that there exists a partition $(\mathcal{I}_z)_{z=1,\dots,Z} \subset 2^{\mathcal{I}}$ of $\mathcal{I}$, into Z (e.g., geographical) subsets, where $2^{\mathcal{I}}$ is the power set of $\mathcal{I}$ with $Z \leqslant |\mathcal{I}|$ obviously. The binary variable vector $u := (u_{jz})_{j\in\mathcal{J},z=1,\dots,Z+1} \in \{0,1\}^{|\mathcal{J}|(Z+1)}$ represents the allocations of demand points to subsets, i.e., u_{jz} implies that demand point j has been assigned to a facility in subset z. Note that we introduce a dummy subset $\mathcal{I}_{Z+1} := \varnothing$ to which we allocate demand points that have not been assigned to any facility. Hence, the subsets $(\mathcal{I}_z)_{z=1,\dots,Z+1}$ form likewise a partition of $\mathcal{I}$.

We let the assignment of demand points to subsets determine the joint distribution of ξ on $\mathbb{R}^{|\mathcal{J}|}$. This gives $(Z+1)^{|\mathcal{J}|}$ potential distributions, one for each possible allocation of demand points to subsets. Observe that this allows us to model a vast array of potential configurations. At one end, we find the original case where all locations are in the same subset and the distribution is not affected by allocation decisions. At the other end, we find the case where $Z = (|\mathcal{I}| + 1)$ and the joint demand distribution depends on the individual facilities to which demand points have been allocated. By acting on the number of subsets Z, one obtains that the distribution depends on the subset of facilities to which demand points are assigned. To gain some practical intuition, one may think of subsets as, e.g., geographical zones or facility

types. Let then $\mathcal{D}$ be the set of all possible joint distributions of demand that may materialize: $|\mathcal{D}| = (Z+1)^{|\mathcal{J}|}$. This gives a probability distribution P_d for each possible allocation decision $d \in \mathcal{D}$. One may for example think of a distribution parametrized by a function of u, i.e., P^θ, with $\theta := f(u)$. Given that u can take only finitely many values, one can enumerate the possible specifications of P^θ as $\mathsf{P}_1, \ldots, \mathsf{P}_{|\mathcal{D}|}$. Let each component δ_d of the binary variable vector $\delta := (\delta_d)_{d\in\mathcal{D}} \in \{0,1\}^{|\mathcal{D}|}$ take value one if distribution d is enforced by the allocation decisions, and be equal to 0 otherwise. Finally, we define a map $\zeta : \mathcal{J} \times \mathcal{D} \to \{1, \ldots, Z+1\}$, which returns the zone to which a demand point j must be assigned in order for the distribution d to accurately describe the uncertain demand. Having this notation in place, the *single-source facility location problem with decision-dependent stochastic demand* can be expressed as follows:

$$\min \sum_{i\in\mathcal{I}} F_i x_i + \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} C_{ij} y_{ij} - \sum_{d\in\mathcal{D}} \delta_d \mathbb{E}_{\mathsf{P}_d} Q(y,\xi) \tag{1a}$$

$$\text{s.t. } y_{ij} \leqslant x_i \qquad \forall i \in \mathcal{I}, j \in \mathcal{J} \tag{1b}$$

$$\sum_{i\in\mathcal{I}} y_{ij} \leqslant 1 \qquad \forall j \in \mathcal{J} \tag{1c}$$

$$\sum_{i\in\mathcal{I}_z} y_{ij} = u_{jz} \qquad \forall j \in \mathcal{J}, z = 1, \ldots, Z \tag{1d}$$

$$\sum_{z=1}^{Z+1} u_{jz} = 1 \qquad \forall j \in \mathcal{J} \tag{1e}$$

$$\sum_{j\in\mathcal{J}} u_{j,\zeta(j,d)} \leqslant |\mathcal{J}| + \delta_d - 1 \qquad \forall d \in \mathcal{D} \tag{1f}$$

$$\sum_{j\in\mathcal{J}} u_{j,\zeta(j,d)} \geqslant |\mathcal{J}|\delta_d \qquad \forall d \in \mathcal{D} \tag{1g}$$

$$x_i \in \{0,1\} \qquad \forall i \in \mathcal{I} \tag{1h}$$

$$y_{ij} \in \{0,1\} \qquad \forall i \in \mathcal{I}, j \in \mathcal{J} \tag{1i}$$

$$u_{jz} \in \{0,1\} \qquad \forall j \in \mathcal{J}, z = 1, \ldots, Z+1 \tag{1j}$$

$$\delta_d \in \{0,1\} \qquad \forall d \in \mathcal{D} \tag{1k}$$

The objective function sums the costs of opening facilities and links to demand points, and accounts for the stochastic revenue made as a consequence of such decisions. Here $\mathbb{E}_{\mathsf{P}_d} Q(y,\xi) = \int Q(y,\xi)\, \mathrm{d}\mathsf{P}_d$ represents the expected second-stage revenue according to distribution P_d. We assume this expectation to be finite for all y and $d \in \mathcal{D}$. The second-stage function $Q(y,\xi)$ will be introduced later.

Constraints (1b) state that a connection can be open only if the corresponding facility is open. Constraints (1c) state that each demand point can be connected to at most one facility. Constraints (1d) set $u_{jz} = 1$ whenever demand point j has been assigned to a facility in set z. Constraints (1e) ensure that each demand point is either assigned to a real subset or to the dummy set $Z+1$. Recall that the assignment of demand points to subsets determines the demand distribution. Here constraints (1f) model the implication $\sum_{j\in\mathcal{J}} u_{j,\zeta(j,d)} = |\mathcal{J}| \implies \delta_d = 1$ while (1g) implement $\delta_d = 1 \implies \sum_{j\in\mathcal{J}} u_{j,\zeta(j,d)} = |\mathcal{J}|$. This double implication ensures that

distribution P_d is enforced if and only if the corresponding allocation of demand points to subsets, according to ζ, holds.

Finally, we represent the second-stage problem as follows:

$$Q(y,\xi) = \max \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} R_{ij} w_{ij} \tag{2a}$$

$$\text{s.t. } w_{ij} \leqslant \xi_j y_{ij} \qquad \forall i \in \mathcal{I}, j \in \mathcal{J} \tag{2b}$$

$$\sum_{j\in\mathcal{J}} w_{ij} \leqslant M_i \qquad \forall i \in \mathcal{I} \tag{2c}$$

$$w_{ij} \geqslant 0 \qquad \forall i \in \mathcal{I}, j \in \mathcal{J} \tag{2d}$$

Here w_{ij} represents the demand of demand point j satisfied by facility i. The objective is to maximize revenues, where $R_{ij} \geqslant 0$ is the per-unit revenue for the demand of demand point j satisfied from facility i. The constraints ensure that demands ξ_j and facility capacities $M_i \geqslant 0$ are not exceeded.

Observe that model (1) is in general nonlinear due to the product in the last term of the objective function. For discrete probability distributions defined with finitely many scenarios, we obtain a bilinear optimization problem due to the product of variables δ_d and w_{ij} in the objective function. These bilinear terms can be linearized exactly using the McCormick method [McCormick, 1976], which yields an MILP problem. Nevertheless, the number of constraints (1f)-(1g) is exponential in $|\mathcal{J}|$.

3 Solution Algorithm

We develop a cutting-plane algorithm that instantiates and extends the L-Shaped method for stochastic programs with decision-dependent uncertainty [Pantuso and Hewitt, 2025]. In addition, we provide a novel, generally-applicable, optimality cut, which is tighter than one developed by Pantuso and Hewitt [2025], but requires dealing with bilinear terms that can be linearized exactly.

In Section 3.1, we first present the L-Shaped method for stochastic programs with decision-dependent uncertainty. Then, in Section 3.2, we introduce the new optimality cuts. In Section 3.3, we introduce a problem-specific valid inequality.

3.1 The L-Shaped Method with Decision-Dependent Distributions

In this section, we extend the method of Pantuso and Hewitt [2025] to our problem. We assume that ξ is a discrete random variable so that distributions P_d, $d \in \mathcal{D}$, are step functions. Let ξ_{ds} and π_{ds} respectively denote the realization of the random variable ξ and its probability under scenario s of distribution P_d and let $\mathcal{S}_d$ be the distribution-specific set of scenarios characterizing P_d.

We first obtain the epigraph reformulation of problem (1) by adding a decision variable μ:

$$\min \sum_{i\in\mathcal{I}} F_i x_i + \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} C_{ij} y_{ij} - \mu \tag{3a}$$

$$\text{s.t. } (1b) to (1k) \tag{3b}$$

$$\mu \leqslant \sum_{d\in\mathcal{D}} \delta_d \mathbb{E}_{\mathsf{P}_d} Q(y,\xi) \tag{3c}$$

$$\mu \geqslant 0 \tag{3d}$$

Observe that, given that the subsets $(\mathcal{I}_z)_{z=1,\dots,Z+1}$ form a partition of $\mathcal{I}$, only one δ_d can take value 1. Furthermore, observe that $Q(y,\xi)$ is always bounded from above under the reasonable assumption that M_i and ξ_j are finite. In addition, since $(0)_{i\in\mathcal{I},j\in\mathcal{J}}$ is always feasible, it follows that $Q(y,\xi) \geqslant 0$ for all y and ξ and that there exists an upper bound U on the second-stage expected revenue given by:

$$U \geqslant \max_{d\in\mathcal{D}} \max_{y\in\{0,1\}^{|\mathcal{I}||\mathcal{J}|}} \mathbb{E}_{\mathsf{P}_d} Q(y,\xi) \ .$$

Proposition 1 provides a closed-form expression for U.

Proposition 1. *Let*

$$U_1 := \sum_{j\in\mathcal{J}} \max_{i\in\mathcal{I}} \left\{ R_{ij} \min\left\{ M_i, \max_{d\in\mathcal{D}} \mathbb{E}_{\mathsf{P}_d}[\xi_j] \right\} \right\}$$

and

$$U_2 := \left(\max_{i\in\mathcal{I},j\in\mathcal{J}} R_{ij} \right) \sum_{i\in\mathcal{I}} M_i$$

Then $U := \min\{U_1, U_2\}$ *satisfies* $U \geqslant \mathbb{E}_{\mathsf{P}_d} Q(y,\xi)$ *for every* $d \in \mathcal{D}$ *and every* y *satisfying* (1c) *and* (1i).

Proof. It suffices to prove the bound for U_1 and for U_2 separately.

Consider U_2. Let w be feasible for $Q(y,\xi)$. Since $w \geqslant 0$ and $R_{ij} \leqslant \max_{i,j} R_{ij}$ for all i,j, constraints (2c) give

$$\sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} R_{ij} w_{ij} \leqslant \left(\max_{i,j} R_{ij} \right) \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} w_{ij} \leqslant \left(\max_{i,j} R_{ij} \right) \sum_{i\in\mathcal{I}} M_i = U_2$$

Hence $Q(y,\xi) \leqslant U_2$ for every y and every realization ξ, and therefore $\mathbb{E}_{\mathsf{P}_d} Q(y,\xi) \leqslant U_2$ for every d.

Consider now U_1. Fix $d \in \mathcal{D}$ and y. The function $Q(y,\cdot)$ is the optimal value of a maximization linear program in which ξ enters only the right-hand side of (2b). It follows that $Q(y,\cdot)$ is concave in ξ. Jensen's inequality gives

$$\mathbb{E}_{\mathsf{P}_d} Q(y,\xi) \leqslant Q\left(y, \mathbb{E}_{\mathsf{P}_d}[\xi]\right)$$

Let w be feasible for $Q(y, \mathbb{E}_{\mathsf{P}_d}[\xi])$. By (2b), $w_{ij} \leqslant \mathbb{E}_{\mathsf{P}_d}[\xi_j] y_{ij}$, and by (2c) together with $w \geqslant 0$, $w_{ij} \leqslant \sum_{j'\in\mathcal{J}} w_{ij'} \leqslant M_i$, so that, for all $i \in \mathcal{I}$ and $j \in \mathcal{J}$ we have

$$w_{ij} \leqslant \min\left\{ M_i, \mathbb{E}_{\mathsf{P}_d}[\xi_j] \right\} y_{ij}$$

Therefore

$$Q\left(y, \mathbb{E}_{\mathsf{P}_d}[\xi]\right) \leqslant \sum_{j\in\mathcal{J}} \sum_{i\in\mathcal{I}} R_{ij} \min\left\{ M_i, \mathbb{E}_{\mathsf{P}_d}[\xi_j] \right\} y_{ij}$$

By (1c) and (1i), for each $j \in \mathcal{J}$ at most one y_{ij} equals one. Hence, the inner sum is bounded by its largest term, that is

$$\sum_{i\in\mathcal{I}} R_{ij} \min\left\{M_i, \mathbb{E}_{\mathsf{P}_d}[\xi_j]\right\} y_{ij} \leqslant \max_{i\in\mathcal{I}} \left\{R_{ij} \min\left\{M_i, \mathbb{E}_{\mathsf{P}_d}[\xi_j]\right\}\right\}$$

Finally, $t \mapsto \min\{M_i, t\}$ is non-decreasing, so $\min\{M_i, \mathbb{E}_{\mathsf{P}_d}[\xi_j]\} \leqslant \min\{M_i, \max_{d'\in\mathcal{D}} \mathbb{E}_{\mathsf{P}_{d'}}[\xi_j]\}$ for every d. Summing over $j \in \mathcal{J}$ yields $\mathbb{E}_{\mathsf{P}_d} Q(y,\xi) \leqslant U_1$. Since d and y were arbitrary the claim follows. □

As a result, we can rewrite (3c) as follows:

$$\mu \leqslant \mathbb{E}_{\mathsf{P}_d} Q(y,\xi) + U(1-\delta_d) \qquad \forall d \in \mathcal{D} \tag{4}$$

Note that the variable δ_d, which takes value 1 when distribution effectively d is enforced, will set the tightest bound on μ and make constraints (4) equivalent to (3c).

Since ξ is discrete, we may then write (4) as follows:

$$\mu \leqslant \sum_{s\in\mathcal{S}_d} \pi_{ds} Q(y,\xi_{ds}) + U(1-\delta_d) \qquad \forall d \in \mathcal{D}\ . \tag{5}$$

Since each $Q(y,\xi_{ds})$ is always feasible and bounded, strong duality implies that the optimal solution to the dual problem is always attained at one of the extreme points of the set:

$$\left\{(\rho,\gamma) \in \mathbb{R}^{|\mathcal{I}||\mathcal{J}|}_{\geqslant 0} \times \mathbb{R}^{|\mathcal{I}|}_{\geqslant 0} \,|\, \gamma_i + \rho_{ij} \geqslant R_{ij}, i \in \mathcal{I}, j \in \mathcal{J}\right\}\ .$$

Let (ρ^k, γ^k) be the k-th extreme point and $\mathcal{R} = \{(\rho^1,\gamma^1), \ldots, (\rho^K,\gamma^K)\}$ be the collection of such extreme points. Hence, for each $d \in \mathcal{D}$ and $s \in \mathcal{S}_d$, we have:

$$Q(y,\xi_{ds}) = \min_{(\rho,\gamma)\in\mathcal{R}} \left\{\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \rho_{ij}\xi_{jds} y_{ij} + \sum_{i\in\mathcal{I}} \gamma_i M_i\right\}\ ,$$

which allows the rewriting of (5) as

$$\mu \leqslant \sum_{s\in\mathcal{S}_d} \pi_{ds} \left(\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \rho^{sd}_{ij}\xi_{jds} y_{ij} + \sum_{i\in\mathcal{I}} \gamma^{sd}_i M_i\right) + U(1-\delta_d),\ \forall d \in \mathcal{D}, (\rho^{sd},\gamma^{sd})_{s\in\mathcal{S}_d} \in \times_{s\in\mathcal{S}_d}\mathcal{R}, \tag{6}$$

where $\times_{s\in\mathcal{S}_d}$ denotes the cartesian product. Note that the last term in (6) models the implication:

$$\delta_d = 1 \implies \mu \leqslant \sum_{s\in\mathcal{S}_d} \pi_{ds} \left(\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \rho^{sd}_{ij}\xi_{jds} y_{ij} + \sum_{i\in\mathcal{I}} \gamma^{sd}_i M_i\right) \forall (\rho^{sd},\gamma^{sd})_{s\in\mathcal{S}_d} \in \times_{s\in\mathcal{S}_d}\mathcal{R}\ .$$

As presented next, the same implication can be modeled using only the variables u. Recall the map $\zeta : \mathcal{J}\times\mathcal{D} \to \{1,\ldots,Z+1\}$, which returns the subset z to which j has to be assigned in order for distribution P_d to apply. Recalling that $|\mathcal{J}| - \sum_{j\in\mathcal{J}} u_{j,\zeta(j,d)} = 0 \implies \sum_{j\in\mathcal{J}} u_{j,\zeta(j,d)} = |\mathcal{J}| \implies \delta_d = 1$, we can rewrite (6) as

$$\mu \leqslant \sum_{s\in\mathcal{S}_d} \pi_{ds} \left(\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \rho^{sd}_{ij}\xi_{jds} y_{ij} + \sum_{i\in\mathcal{I}} \gamma^{sd}_i M_i\right) + U\left(|\mathcal{J}| - \sum_{j\in\mathcal{J}} u_{j,\zeta(j,d)}\right),\ \forall d \in \mathcal{D}, (\rho^{sd},\gamma^{sd})_{s\in\mathcal{S}_d} \in \times_{s\in\mathcal{S}_d}\mathcal{R}, \tag{7}$$

where we have:

$$|\mathcal{J}| - \sum_{j\in\mathcal{J}} u_{j,\zeta(j,d)} = 0 \implies \mu \leqslant \sum_{s\in\mathcal{S}_d} \pi_{ds} \left(\sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \rho_{ij}^{sd} \xi_{jds} y_{ij} + \sum_{i\in\mathcal{I}} \gamma_i^{sd} M_i \right) \forall (\rho^{sd}, \gamma^{sd})_{s\in\mathcal{S}_d} \in \times_{s\in\mathcal{S}_d} \mathcal{R}.$$

This effectively makes the variables δ redundant. Hence we arrive at the following reformulation of (1), which we will refer to as the Master Problem (MP):

$$\min \left\{ \sum_{i\in\mathcal{I}} F_i x_i + \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} C_{ij} y_{ij} - \mu | (1b), (1c), (1d), (1e), (1h), (1i), (1j), (7) \right\} . \tag{MP}$$

Observe that, compared to (1), (MP) does not include constraints (1f) and (1g) and the δ_d variables (1k), which are specified for all elements in $\mathcal{D}$. On the other hand, it includes constraints (7), which require enumerating the extreme points of the dual of problem $Q(y, \xi_{ds})$ for each d.

We propose to solve (MP) by means of a cutting-plane algorithm, which relaxes constraints (7) and adds cuts when found violated. Hence, at each iteration we solve the following *Relaxed Master Problem* (RMP):

$$\min \left\{ \sum_{i\in\mathcal{I}} F_i x_i + \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} C_{ij} y_{ij} - \mu | (1b), (1c), (1d), (1e), (1h), (1i), (1j) \right\} . \tag{RMP}$$

Observe that (RMP), as displayed, contains no constraint bounding μ from above. This would result in an unbounded problem when no optimality cut is present. We therefore impose $\mu \leqslant U$, which is valid by definition of U and never binding at an optimal solution of (MP).

The procedure is sketched in Algorithm 1. At every iteration v, (RMP) is solved with the currently available constraints (7) (initially none). Given a solution (x^v, y^v, u^v, μ^v) for (RMP), we identify the distribution P_{d^v} enforced by u^v, that is by the allocation of demand points to subsets of $\mathcal{I}$. We compute next the expected second-stage revenue associated to that probability distribution. Observe at line 6 of the pseudo-code that

$$\mathbb{E}_{\mathsf{P}_{d^v}}[Q(y^v, \xi)] = \sum_{s\in\mathcal{S}_d} \pi_{d^v s} Q(y^v, \xi_{d^v s}) = \sum_{s\in\mathcal{S}_{d^v}} \pi_{d^v s} \left(\sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \rho_{ij}^{sd^v v} \xi_{jd^v s} y_{ij}^v + \sum_{i\in\mathcal{I}} \gamma_i^{sd^v v} M_i \right) ,$$

where $\rho_{ij}^{sd^v v}$ and $\gamma_i^{sd^v v}$ are the dual solution to $Q(y^v, \xi_{d^v s})$. Hence, computing the expectation amounts to solving the second-stage problems only for the scenarios of distribution d^v. If μ^v holds the correct value of the expectation, the algorithm has converged to an optimal solution. Otherwise, we generate a constraint (7) for $d = d^v$. Particularly, if $\mu^v > \mathbb{E}_{\mathsf{P}_{d^v}}[Q(y^v, \xi)]$, we identify a violated inequality (7) as:

$$\mu \leqslant \sum_{s\in\mathcal{S}_{d^v}} \pi_{d^v s} \left(\sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \rho_{ij}^{sd^v v} \xi_{jd^v s} y_{ij} + \sum_{i\in\mathcal{I}} \gamma_i^{sd^v v} M_i \right) + U \left(|\mathcal{J}| - \sum_{j\in\mathcal{J}} u_{j,\zeta(j,d^v)} \right) . \tag{OC}$$

Observe that (x^v, y^v, u^v, μ^v) violates (OC) by effect of having $|\mathcal{J}| = \sum_{j\in\mathcal{J}} u^v_{j,\zeta(j,d^v)}$. Hence, we add (OC) to (RMP) as an *optimality cut* that renders (x^v, y^v, u^v, μ^v) infeasible in the subsequent iterations. Observe also how (OC) is redundant for distributions other than d^v.

Proposition 2. *Algorithm 1 converges in a finite number of iterations.*

Proof. As we have shown, (MP) is a reformulation of (1). Recovering all (finitely many) inequalities (7) is a sufficient condition for convergence. Therefore, the algorithm terminates in a finite number of iterations. □

```
solved←false;
v ← 0;
while not solved do
    Solve (RMP). Let (x^v, y^v, u^v, μ^v) be its optimal solution;
    Identify the distribution P_{d^v}, d^v ∈ 𝒟, enforced by allocation decisions u^v;
    Compute 𝔼_{P_{d^v}}[Q(y^v, ξ)];
    if μ^v > 𝔼_{P_{d^v}}[Q(y^v, ξ)] then
        Identify a violated constraint (7) and add it to (RMP);
    end
    else
        solved←true;
    end
    v ← v + 1;
end
```

Algorithm 1: Decision-Dependent L-Shaped Method for (1).

We stress that, in practical implementations, Algorithm 1 is embedded in a Branch & Cut framework that solves (RMP). That is, Algorithm 1 is initiated upon reaching Branch & Cut nodes which produce an integer feasible solution to (RMP). The optimality cuts obtained are, however, globally valid and can hence be broadcast to all open nodes of the Branch & Cut tree.

3.2 Tighter Optimality Cut

In this section, we derive a new optimality cut. Throughout, consider a generic iteration v at which the cut is generated. Denote then d^v the distribution induced by the solution to (RMP) and $(\rho^{sd^vv}, \gamma^{sd^vv})_{s\in\mathcal{S}_{d^v}}$ the dual solutions to $Q(y^v, \xi_s)$, $s \in \mathcal{S}_{d^v}$. With these elements we define

$$a^v_{ij} := \sum_{s\in\mathcal{S}_{d^v}} \pi_{d^vs}\rho^{sd^vv}_{ij}\xi_{jd^vs} \quad \text{and} \quad c_v := \sum_{s\in\mathcal{S}_{d^v}} \pi_{d^vs}\sum_{i\in\mathcal{I}}\gamma^{sd^vv}_i M_i \tag{8}$$

so that the right-hand side of (OC) (excluding the last term) reads

$$V^v(y) := c_v + \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} a^v_{ij}y_{ij} \tag{9}$$

Observe that the coefficients (8) are nonnegative, because ρ^{sd^vv} and γ^{sd^vv} belong to $\mathcal{R} \subset \mathbb{R}^{|\mathcal{I}||\mathcal{J}|}_{\geqslant 0} \times \mathbb{R}^{|\mathcal{I}|}_{\geqslant 0}$, while probabilities, demands and capacities are nonnegative as well. Observe further that

each demand point is assigned to at most one facility. It follows that $V^v(y)$ ranges in $[c_v, \bar{B}_v]$, where

$$\bar{B}_v := c_v + \sum_{j\in\mathcal{J}} \max_{i\in\mathcal{I}} a^v_{ij} \tag{10}$$

We finally let $K_d(u) := |\mathcal{J}| - \sum_{j\in\mathcal{J}} u_{j,\zeta(j,d)}$ denote the number of demand points whose allocation is different from the one that enforces P_d. Then, (OC) reads $\mu \leqslant V^v(y) + K_{d^v}(u)U$.

Proposition 3. *Let $B_v := \max\{U, \bar{B}_v\}$. The optimality cut*

$$\mu \leqslant \sum_{s\in\mathcal{S}_{d^v}} \pi_{d^v s} \left(\sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \rho^{sd^v v}_{ij} \xi_{jd^v s} y_{ij} + \sum_{i\in\mathcal{I}} \gamma^{sd^v v}_i M_i \right) + B_v \left(|\mathcal{J}| - \sum_{j\in\mathcal{J}} u_{j,\zeta(j,d^v)} \right) \tag{11a}$$

$$- \sum_{s\in\mathcal{S}_{d^v}} \pi_{d^v s} \left(\sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \rho^{sd^v v}_{ij} \xi_{jd^v s} y_{ij} + \sum_{i\in\mathcal{I}} \gamma^{sd^v v}_i M_i \right) \left(|\mathcal{J}| - \sum_{j\in\mathcal{J}} u_{j,\zeta(j,d^v)} \right) \tag{11b}$$

is valid for (MP). In addition, it is equivalent to (OC) when $K_{d^v}(u) = 0$, while for $K_{d^v}(u) \geqslant 1$, it is strictly tighter than (OC) if $V^v(y) > B_v - U$.

Proof. We start by showing validity. Using (9), cut (11) reads $\mu \leqslant V^v(y) + K_{d^v}(u)\,(B_v - V^v(y))$. Let (x, y, u, μ) be feasible for (MP). Let $d(u)$ be the distribution enforced by u. If $d(u) = d^v$ then $K_{d^v}(u) = 0$. Then, the right-hand side equals $V^v(y)$. Since $(\rho^{sd^v v}, \gamma^{sd^v v}) \in \mathcal{R}$ is feasible for the dual of $Q(y, \xi_{d^v s})$ for every y, weak duality gives $V^v(y) \geqslant \mathbb{E}_{\mathsf{P}_{d^v}} Q(y, \xi) \geqslant \mu$. Here, the last inequality follows from (3c). Othwerwise, if $d(u) \neq d^v$ let $K := K_{d^v}(u) \geqslant 1$. Because $a^v \geqslant 0$ and each demand point is assigned to at most one facility, $V^v(y) \leqslant \bar{B}_v \leqslant B_v$. Rewriting the right-hand side as $B_v + (K-1)\big(B_v - V^v(y)\big)$, we see that both terms are nonnegative increments over B_v. It follows that

$$V^v(y) + K\,(B_v - V^v(y)) \geqslant B_v \geqslant U \geqslant \mathbb{E}_{\mathsf{P}_{d(u)}} Q(y, \xi) \geqslant \mu$$

This confirms that cut (11) is valid.

For the comparison with (OC), it is easy to see that subtracting the two right-hand sides gives $K_{d^v}(u)\,(B_v - U - V^v(y))$, which is zero when $K_{d^v}(u) = 0$ and negative exactly when $V^v(y) > B_v - U$. □

Remark 1. *Note that (OC) and (11) are both valid but are not ordered. It follows that separating both in Algorithm 1 provides, jointly, a tighter cut that either alone.*

Note that the optimality cut (11) introduces bilinear terms. By expanding the products in (11b), we successively obtain

$$\sum_{s\in\mathcal{S}_{d^v}} \pi_{d^v s} \left(\sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \rho^{sd^v v}_{ij} \xi_{jd^v s} y_{ij} + \sum_{i\in\mathcal{I}} \gamma^{sd^v v}_i M_i \right) \left(|\mathcal{J}| - \sum_{j\in\mathcal{J}} u_{j,\zeta(j,d^v)} \right) \tag{12}$$

$$= \left(\sum_{s\in\mathcal{S}_{d^v}} \pi_{d^v s} \left(\sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \rho^{sd^v v}_{ij} \xi_{jd^v s} y_{ij} \right) + \sum_{s\in\mathcal{S}_{d^v}} \pi_{d^v s} \left(\sum_{i\in\mathcal{I}} \gamma^{sd^v v}_i M_i \right) \right) \left(|\mathcal{J}| - \sum_{j\in\mathcal{J}} u_{j,\zeta(j,d^v)} \right) \tag{13}$$

$$= \sum_{s\in\mathcal{S}_{d^v}} \pi_{d^v s} \left(\sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \rho_{ij}^{sd^v v} \xi_{jd^v s} y_{ij} \right) \left(|\mathcal{J}| - \sum_{j\in\mathcal{J}} u_{j,\zeta(j,d^v)} \right) \tag{14}$$

$$+ \sum_{s\in\mathcal{S}_{d^v}} \pi_{d^v s} \left(\sum_{i\in\mathcal{I}} \gamma_i^{sd^v v} M_i \right) \left(|\mathcal{J}| - \sum_{j\in\mathcal{J}} u_{j,\zeta(j,d^v)} \right) \tag{15}$$

$$= \sum_{s\in\mathcal{S}_{d^v}} \pi_{d^v s} \left(\sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \rho_{ij}^{sd^v v} \xi_{jd^v s} |\mathcal{J}| y_{ij} \right) \tag{16}$$

$$- \sum_{s\in\mathcal{S}_{d^v}} \pi_{d^v s} \left(\sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \rho_{ij}^{sd^v v} \xi_{jd^v s} y_{ij} \left(\sum_{j_2\in\mathcal{J}} u_{j_2,\zeta(j,d^v)} \right) \right) \tag{17}$$

$$+ \sum_{s\in\mathcal{S}_{d^v}} \pi_{d^v s} \left(\sum_{i\in\mathcal{I}} \gamma_i^{sd^v v} M_i \right) \left(|\mathcal{J}| - \sum_{j\in\mathcal{J}} u_{j,\zeta(j,d^v)} \right) \tag{18}$$

$$= \sum_{s\in\mathcal{S}_{d^v}} \pi_{d^v s} \left(\sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \rho_{ij}^{sd^v v} \xi_{jd^v s} |\mathcal{J}| y_{ij} \right) \tag{19}$$

$$- \sum_{s\in\mathcal{S}_{d^v}} \pi_{d^v s} \left(\sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \rho_{ij}^{sd^v v} \xi_{jd^v s} \sum_{j_2\in\mathcal{J}} y_{ij} u_{j_2,\zeta(j,d^v)} \right) \tag{20}$$

$$+ \sum_{s\in\mathcal{S}_{d^v}} \pi_{d^v s} \left(\sum_{i\in\mathcal{I}} \gamma_i^{sd^v v} M_i \right) \left(|\mathcal{J}| - \sum_{j\in\mathcal{J}} u_{j,\zeta(j,d^v)} \right) \tag{21}$$

$$= \sum_{s\in\mathcal{S}_{d^v}} \pi_{d^v s} \left(\sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \rho_{ij}^{sd^v v} \xi_{jd^v s} |\mathcal{J}| y_{ij} \right) - \sum_{s\in\mathcal{S}_{d^v}} \pi_{d^v s} \left(\sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \rho_{ij}^{sd^v v} \xi_{jd^v s} \sum_{j_2\in\mathcal{J}} y_{ij} u_{j_2,\zeta(j,d^v)} \right) \tag{22}$$

$$+ \sum_{s\in\mathcal{S}_{d^v}} \pi_{d^v s} \left(\sum_{i\in\mathcal{I}} \gamma_i^{sd^v v} M_i |\mathcal{J}| \right) - \sum_{s\in\mathcal{S}_{d^v}} \pi_{d^v s} \left(\sum_{i\in\mathcal{I}} \gamma_i^{sd^v v} M_i \sum_{j\in\mathcal{J}} u_{j,\zeta(j,d^v)} \right). \tag{23}$$

We can now rewrite the optimality cut as follows:

$$\mu \leqslant \sum_{s\in\mathcal{S}_{d^v}} \pi_{d^v s} \left(\sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \rho_{ij}^{sd^v v} \xi_{jd^v s} y_{ij} + \sum_{i\in\mathcal{I}} \gamma_i^{sd^v v} M_i \right) + B_v \left(|\mathcal{J}| - \sum_{j\in\mathcal{J}} u_{j,\zeta(j,d^v)} \right) \tag{24}$$

$$- \sum_{s\in\mathcal{S}_{d^v}} \pi_{d^v s} \left(\sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \rho_{ij}^{sd^v v} \xi_{jd^v s} |\mathcal{J}| y_{ij} \right) + \sum_{s\in\mathcal{S}_{d^v}} \pi_{d^v s} \left(\sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \rho_{ij}^{sd^v v} \xi_{jd^v s} \sum_{j_2\in\mathcal{J}} y_{ij} u_{j_2,\zeta(j,d^v)} \right) \tag{25}$$

$$- \sum_{s\in\mathcal{S}_{d^v}} \pi_{d^v s} \left(\sum_{i\in\mathcal{I}} \gamma_i^{sd^v v} M_i |\mathcal{J}| \right) + \sum_{s\in\mathcal{S}_{d^v}} \pi_{d^v s} \left(\sum_{i\in\mathcal{I}} \gamma_i^{sd^v v} M_i \sum_{j\in\mathcal{J}} u_{j,\zeta(j,d^v)} \right) \tag{26}$$

The bilinear terms have now been isolated in the second term of (25) and involve only products of binary variables. They can be linearized using the Fortet linearization approach [Fortet, 1960] described next. We replace the products $y_{ij_1} u_{j_2,z}$ by adding to (RMP) new variables $\ell_{ij_1j_2z}$, for $i \in \mathcal{I}$, $j_1, j_2 \in \mathcal{J}$, $z \in \mathcal{Z}$. The new variables are constrained by the Fortet inequalities (27):

$$\ell_{ij_1j_2z} \leqslant y_{ij_1} \qquad i \in \mathcal{I}, j_1, j_2 \in \mathcal{J}, z \in \mathcal{Z} \tag{27a}$$

$$\ell_{ij_1j_2z} \leqslant u_{j_2z} \qquad i \in \mathcal{I}, j_1, j_2 \in \mathcal{J}, z \in \mathcal{Z} \tag{27b}$$

$$\ell_{ij_1j_2z} + 1 \geqslant y_{ij_1} + u_{j_2z} \qquad i \in \mathcal{I}, j_1, j_2 \in \mathcal{J}, z \in \mathcal{Z} \tag{27c}$$

$$\ell_{ij_1j_2z} + 1 \geqslant 0 \qquad i \in \mathcal{I}, j_1, j_2 \in \mathcal{J}, z \in \mathcal{Z} \tag{27d}$$

Constraints (27d) are always satisfied when the ℓ variables are nonnegative. The new linear optimality cut can then be reformulated as:

$$\begin{aligned}\mu \leqslant & \sum_{s\in\mathcal{S}_{d^v}} \pi_{d^v s}\left(\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \rho_{ij}^{sd^v v}\xi_{jd^v s}y_{ij} + \sum_{i\in\mathcal{I}}\gamma_i^{sd^v v}M_i\right) + B_v\left(|\mathcal{J}| - \sum_{j\in\mathcal{J}} u_{j,\zeta(j,d^v)}\right)\\
& - \sum_{s\in\mathcal{S}_{d^v}} \pi_{d^v s}\left(\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \rho_{ij}^{sd^v v}\xi_{jd^v s}|\mathcal{J}|y_{ij}\right)\\
& + \sum_{s\in\mathcal{S}_{d^v}} \pi_{d^v s}\left(\sum_{i\in\mathcal{I}}\sum_{j_1\in\mathcal{J}}\sum_{j_2\in\mathcal{J}} \rho_{ij_1}^{sd^v v}\xi_{j_1 d^v s}\ell_{ij_1 j_2\zeta(j_2,d^v)}\right)\\
& - \sum_{s\in\mathcal{S}_{d^v}} \pi_{d^v s}\left(\sum_{i\in\mathcal{I}}\gamma_i^{sd^v v}M_i|\mathcal{J}|\right) + \sum_{s\in\mathcal{S}_{d^v}} \pi_{d^v s}\left(\sum_{i\in\mathcal{I}}\gamma_i^{sd^v v}M_i\sum_{j\in\mathcal{J}} u_{j,\zeta(j,d^v)}\right).\end{aligned} \tag{28}$$

Observe that this reformulation may require a substantial number of linearization variables and constraints (27). For this reason, we have identified different ways to introduce them. The three ways differ in which of the Fortet inequalities (27) are imposed on (RMP) from the outset and which are left to be satisfied otherwise. One point is common to all three and constrains the choice. The variables $\ell_{ij_1j_2z}$ appear in (28) with a nonnegative coefficient, so the cut is valid only where $\ell_{ij_1j_2z} \geqslant y_{ij_1}u_{j_2z}$. Constraints (27c) are the ones enforcing that inequality, and every version must therefore guarantee them.

The first way, which we refer to as **version 1** (V1), consists of adding the linearization variables ℓ to (RMP) as continuous variables in $[0,1]$. In addition, we add to (RMP) only the linearization constraints (27a) and (27b). The objective function will ensure satisfaction of constraints (27c) at the optimal solution. That is, when $y_{ij_1} = u_{j_2z} = 1$, the objective will automatically set $\ell_{ij_1j_2z}$ to 1 since it appears with a nonnegative coefficient in (OC) (see Algorithm 2 for the dual solution). Because $\ell_{ij_1j_2z}$ is bounded above by both y_{ij_1} and u_{j_2z}, this upper bound coincides with the product whenever the latter are integer, so (27c) holds at every integer solution and the cut is valid.

The second way, which we refer to as **version 2** (V2), consists of adding the linearization variables ℓ upfront to (RMP), declared as binary, together with the linearization constraints (27c) only. The remaining linearization constraints are added dynamically when found violated. That is, upon reaching integer feasible solutions in the Branch & Cut that solves (RMP), we check whether any of them is violated and we add them as feasibility cuts. To reduce the number of such constraints we separate (29a) in place of (27a) and (27b), which is possible because the variables ℓ are binary.

$$2\ell_{ij_1j_2z} \leqslant y_{ij_1} + u_{j_2z} \qquad i\in\mathcal{I}, j_1, j_2 \in \mathcal{J}, z\in\mathcal{Z}. \tag{29a}$$

Finally, the third way, which we refer to as **version 3** (V3), consists of adding the linearization variables ℓ as binary variables and then adding only (29a) to (RMP) upfront. In this case constraints (27c) are guaranteed at optimality and constraints (27d) are always satisfied. This approach significantly reduces the number of required constraints at the expense of declaring the ℓ variables binary. Here (29a) is present at every node, so when we have $y_{ij_1} = u_{j_2z} = 1$

the variables $\ell_{ij_1j_2z}$ are free up to one and the objective sets them to one, which is the value of the product. Thus, V3 imposes the upper Fortet direction and obtains the lower one from the objective.

To summarise, V1 imposes (27a) and (27b) and obtains (27c) from the objective. V2 imposes (27c) and separates (29a). V3 imposes (29a) and obtains (27c) from the objective. All three carry the same $|\mathcal{I}||\mathcal{J}|^2(Z+1)$ variables ℓ, and, as reported in Appendix A, comparable numbers of constraints. The main difference between the three versions are the inequalities imposed and those left to be recovered.

3.3 Valid Inequality

In this section, we propose a valid inequality that is not distribution-specific and speeds up the convergence of the method.

Proposition 4. *For all $d \in \mathcal{D}$ and for all y, we have*

$$\mathbb{E}_{P_d}[Q(y,\xi)] \leqslant \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} R_{ij} \max_{d\in\mathcal{D}} \min\{M_i, \mathbb{E}_{P_d}[\xi_j]\} y_{ij} \ ,$$

where $\mathsf{P}_d[\xi_j]$ is the marginal expectation of ξ_j according to distribution $d \in \mathcal{D}$.

Proof. Consider any arbitrary y and ξ. Define $w := (w_{ij})_{i\in\mathcal{I},j\in\mathcal{J}}$. The corresponding second-stage problem is:

$$Q(y,\xi) = \max_{w\geqslant 0}\left\{\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} R_{ij}w_{ij}(\xi) | w_{ij} \leqslant \xi_j y_{ij}, i\in\mathcal{I}, j\in\mathcal{J}, \sum_{j\in\mathcal{J}} w_{ij} \leqslant M_i, i\in\mathcal{I}\right\}.$$

Observe that $Q(y,\xi)$ is concave in y and ξ. By Jensen's inequality it follows that, given y, for all $d \in \mathcal{D}$:

$$\begin{aligned}\mathbb{E}_{P_d}[Q(y,\xi)] \leqslant & Q(y, \mathbb{E}_{P_d}[\xi]) \\ = & \max_{w\geqslant 0}\left\{\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} R_{ij}w_{ij}(\xi) | w_{ij} \leqslant \mathbb{E}_{P_d}[\xi_j] y_{ij}, i\in\mathcal{I}, j\in\mathcal{J}, \sum_{j\in\mathcal{J}} w_{ij} \leqslant M_i, i\in\mathcal{I}\right\} \ .\end{aligned}$$

In turn, relaxing the capacity constraints, we obtain:

$$\begin{aligned}& \max_{w\geqslant 0}\left\{\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} R_{ij}w_{ij}(\xi) | w_{ij} \leqslant \mathbb{E}_{P_d}[\xi_j] y_{ij}, i\in\mathcal{I}, j\in\mathcal{J}, \sum_{j\in\mathcal{J}} w_{ij} \leqslant M_i, i\in\mathcal{I}\right\} \\ \leqslant & \max_{w\geqslant 0}\left\{\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} R_{ij}w_{ij}(\xi) | w_{ij} \leqslant \min\{M_i, \mathbb{E}_{P_d}[\xi_j]\} y_{ij}, i\in\mathcal{I}, j\in\mathcal{J}\right\} \\ \leqslant & \max_{w\geqslant 0}\left\{\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} R_{ij}w_{ij}(\xi) | w_{ij} \leqslant \max_{d\in\mathcal{D}}\min\{M_i, \mathbb{E}_{P_d}[\xi_j]\} y_{ij}, i\in\mathcal{I}, j\in\mathcal{J}\right\} \\ = & \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} R_{ij} \max_{d\in\mathcal{D}}\min\{M_i, \mathbb{E}_{P_d}[\xi_j]\} y_{ij} \ .\end{aligned}$$

□

By effect of Proposition 4, the following inequality is valid for all probability distributions

$$\mu \leqslant \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} R_{ij} \max_{d\in\mathcal{D}} \min\{M_i, \mathbb{E}_{P_d}[\xi_j]\} y_{ij} \tag{VI}$$

and can be inserted to strengthen (RMP).

4 Numerical Experiments

We perform numerical experiments on the instances of the problem presented by Laporte et al. [1994]. The instances have been extended in such a way to account for decision-dependent uncertain demand. Following the recipe proposed by Laporte et al. [1994], we randomly generate $|\mathcal{J}|$ demand points with uniform probability in a square of area $[0, 100] \times [0, 100]$. Demands ξ_j are assumed to follow truncated normal distributions parametrized by location decisions as shown later. Each demand point $j \in \mathcal{J}$ is assigned a *base mean* μ_j randomly drawn uniformly in the interval $[10, 50]$ and *base standard deviation* satisfying a ratio σ_j/μ_j randomly drawn uniformly in $[0.05, 0.35]$. As we will see, base mean and standard deviation will form the basis for decision-dependent distributions. Three potential facility locations are then generated in the 10×10 square centered around each of the three demand points with the largest μ_j. The remaining $|\mathcal{I}| - 3$ locations are randomly generated in the 60×60 square centered around the point $(50, 50)$.

We assume that the marginal distribution of ξ_j is fully determined by the facility assigned to the demand point j. To mimic a phenomenon by which demand is positively correlated with proximity to the assigned facility, we proceed as follows. We give subsets $\mathcal{I}_z$, $z = 1, \ldots, Z$ the interpretation of geographical zones. Geographical zones are created by partitioning facility locations into Z clusters according to their coordinates. We use the `k-means++` algorithm, see Arthur and Vassilvitskii [2007].

For each demand point, we rank zones $\mathcal{I}_z$, $z = 1, \ldots, Z$ in non-decreasing order of the distance between the demand point and the centroid of the zone. We let $n(z, j)$ be the rank of zone z in terms of distance from demand point j. Hence, $n(z, j) = 1$ indicates that zone z is the closest to demand point j, while $n(z, j) = Z$ is the furthest and $n(z, j) = Z + 1$ applies to the dummy zone. We assume that the demand at point j follows a normal distribution truncated to the left at 0 with mean and standard deviation depending on the zone to which the demand point is assigned in a given distribution. Particularly, we set the mean demand for demand point j under distribution P_d to:

$$\mu_{jd} = \mu_j(1 + \alpha_j^{n(\zeta(j,d),j)})$$

with $\alpha_j \in [0, 1)$. In all the experiments we set $\alpha_j = 0.5$ for every demand point. That is, the mean under distribution P_d is equal to the base-level mean μ_j, increased by $100\alpha_j^{n(\zeta(j,d),j)}\%$. Observe that the exponent $n(\zeta(j, d), j)$ becomes larger the furthest zone $\zeta(j, d)$ is from j. Hence, the demand increase becomes exponentially smaller as customer j is assigned to more distant zones.

In a similar way, we assume that the standard deviation of ξ_j under distribution P_d is

$$\sigma_{jd} = \sigma_j(1 - \beta_j^{n(\zeta(j,d),j)})$$

with $\beta_j \in [0, 1)$, set to $\beta_j = 0.4$ for every demand point in all the experiments. That is, the standard deviation under distribution P_d is equal to the base-level standard deviation σ_j reduced by $100\beta_j^{n(\zeta(j,d),j)}\%$. In this setting, the exponent $n(\zeta(j,d),j)$ increases as the assigned zone $\zeta(j,d)$ becomes farther from j. Consequently, the reduction in the standard deviation decreases exponentially with the distance between j and its assigned zone.

For each distribution, we generate an empirical distribution by sampling fifty demand realizations (scenarios) of ξ from the underlying truncated normal distributions by inversion.

Remark 2. *Proposition 1 and Proposition 4 hold for any probability measure, and in particular for the empirical measure induced by the sampled scenarios. The expectations $\mathbb{E}_{\mathsf{P}_d}[\xi_j]$ are accordingly computed as sample means over the scenarios that are actually generated, so that U and (VI) apply to the specific problem instance that is solved.*

Consistently with Laporte et al. [1994], we generate test instances that vary in the number of facilities, demand points and zones as well as in the parameters of the problem. The instance sizes considered are $(|\mathcal{I}|, |\mathcal{J}|, Z) \in \{6, 8\} \times \{10, 15, 20\} \times \{3, 4, 5\}$. The parameters are also set according to Laporte et al. [1994] as follows. We set the net revenue per unit of demand of j satisfied by facility i, $R_{ij} = 400$ for all i, j. We set the capacity of facility i as $M_i = M|\mathcal{J}|$ for all i, with $M = 15$. We set the fixed opening cost for facility i as $F_i = O|\mathcal{J}|$ for all i, with $O = 500$. We set the fixed cost for opening link $i - j$ as $C_{ij} = D\Delta_{i,j}$ with $D = 270$ and $\Delta_{i,j}$ being the Euclidean distance between i and j. Finally, since instances are randomly generated, we generate ten random instances for every instance size, which yields a total of $10 \times 18 = 180$ instances. We remark that decision-dependent uncertainty entails working with $(Z + 1)^{|\mathcal{J}|}$ distributions in each instance. In our case, this number ranges from 4^{10} to 6^{20}.

4.1 Results

Algorithm 1 has been implemented in Java using Cplex 22.1.1 libraries with optimality cuts added only to integer nodes in the Branch & Bound tree. Subproblems $Q(y, \xi)$ are solved using a specialized procedure proposed by Laporte et al. [1994] and reported in Appendix B, which returns their primal and dual solutions directly, without invoking a linear programming solver. All tests have been run on a machine equipped with two CPUs Intel(R) Xeon(R) Gold 6230 2.10GHz. Algorithm 1 and all its variants run on a single thread irrespective of the resources made available, and no parallelism is exploited. All tests have been run with Cplex's default settings, including the optimality tolerance of 10^{-4}. The only parameter we set was the time limit, which was set to 1800 seconds in all tests. The times we report are wall-clock times measured around the call to the solver. The optimality gap is calculated as $|\texttt{bestbound} - \texttt{bestinteger}|/(10^{-10} + |\texttt{bestinteger}|)$.

The extensive linearized form of the problem has likewise been solved with Cplex 22.1.1. Nevertheless, given the exponential size of the problem the solver could only solve very small instances ($4-5$ customers). Therefore, no result will be presented for the linearized formulation.

Table 1: Results with and without (VI) with varying number of facilities, demand points, and zones.

			Without (VI)			With (VI)		
$\|\mathcal{I}\|$	$\|\mathcal{J}\|$	Z	`avgGap`[%]	$< 0.5\%$	`avgTime`[sec.]	`avgGap`[%]	$< 0.5\%$	`avgTime`[sec.]
6	10	3	45.64	0%	`lim`	0.00	100%	0.8
6	10	4	57.45	0%	`lim`	0.00	100%	1.0
6	10	5	58.98	0%	`lim`	0.00	100%	2.0
8	10	3	53.90	0%	`lim`	0.00	100%	0.6
8	10	4	59.56	0%	`lim`	0.00	100%	1.1
8	10	5	70.73	0%	`lim`	0.00	100%	1.9

We first analyze the impact of the valid inequality (VI) by describing its effect on some of the small instances. In particular, we used the decomposition algorithm with optimality cuts (OC), with and without (VI). In Table 1 we report the following statistics:

- The average optimality gap (column `avgGap`) as a percentage. This is calculated as $100 * |\texttt{bestbound} - \texttt{bestinteger}|/(10^{-10} + |\texttt{bestinteger}|)$.
- The percentage of instances solved to a gap smaller than 0.5% (column $< 0.5\%$).
- The average solution time (column `avgTime`). The keyword `lim` indicates that every run of the corresponding group terminated at the allocated time limit (1800 seconds).

From Table 1 it is evident that the valid inequality (VI) has a crucial impact on the performance of the algorithm. Without (VI), every one of the sixty runs terminates at the allocated time limit without proving optimality. The corresponding average gaps are between 45.64% and 70.73%. It follows that the bound on μ, which now relies only on the optimality cuts, remains weak and the Branch & Cut has difficulties pruning. With (VI) every instance is solved to optimality in less than two seconds on average.

Part of the explanation behind the impact of (VI) is as follows. In this problem, the space of feasible solutions $\left\{(x_i)_{i\in\mathcal{I}} \in \{0,1\}^{|\mathcal{I}|}, (y_{ij})_{i\in\mathcal{I},j\in\mathcal{J}} \in \{0,1\}^{|\mathcal{I}|\times|\mathcal{J}|}\right\}$ is essentially partitioned into $|\mathcal{D}|$ subsets. Each subset contains the solutions (x, y) that correspond to a specific allocation of demand points to subsets of facilities. Thus, each subset generates a different distribution of the demand. The optimality cuts (OC) reconstruct the expected second-stage revenue function in each of these subsets. The number of subsets is, however, exponentially large in $|\mathcal{J}|$ (recall that in Table 1 we use $|\mathcal{J}| = 10$). Therefore, the algorithm needs to generate a number of cuts sufficient to reconstruct a very large number of expectations. The main effect of (VI) is that it provides non-trivial decision-independent bounds that are valid across multiple subsets of the feasible space. These, in turn, simplify bounding and pruning in the Branch & Cut algorithm. For this reason, all the results presented in what follows are obtained using (VI).

We now compare DD-LS, that is the algorithm with the original optimality cuts (OC), with the three versions V1–V3 of the algorithm that, additionally, separate the new cut (28). Recall

Table 2: Comparison of DD-LS with versions V1–V3 of the algorithm using the new optimality cut (28) together with (OC), by number of demand points. Each cell aggregates the 60 runs obtained by combining $|\mathcal{I}| \in \{6, 8\}$, $Z \in \{3, 4, 5\}$ and the ten instances generated for each size.

	$\lvert\mathcal{J}\rvert$	`solved`	`avgTime`[sec]	`avgGap`[%]
DD-LS	10	52/60	1.3	0.00
	15	31/60	12.2	0.00
	20	29/60	167.9	0.27
V1	10	52/60	6.3	0.00
	15	37/60	146.5	0.12
	20	25/60	501.9	1.01
V2	10	52/60	11.9	0.00
	15	39/60	158.9	0.09
	20	27/60	544.4	1.20
V3	10	48/60	74.8	0.12
	15	19/60	396.9	0.74
	20	20/60	891.1	2.91

from Section 3.2 that the two inequalities are both valid and that neither dominates the other, so that the versions V1–V3 separate both at every round. The comparison is therefore between the original cut alone and the combination of the two. The results are provided in Table 2. They are grouped by number of demand points, which is the main driver both of the number of distributions, $(Z+1)^{|\mathcal{J}|}$, and of the number of linearization variables $\ell_{ij_1j_2z}$. Each cell of Table 2 aggregates the 60 runs obtained by combining $|\mathcal{I}| \in \{6, 8\}$, $Z \in \{3, 4, 5\}$ and the ten instances generated for each size.

On the instances with 15 demand points the combination is clearly stronger than the original cut alone: V2 closes 39 instances out of 60 and V1 closes 37, against 31 for DD-LS. Over the whole test bed V2 solves 118 instances out of 180 and V1 solves 114, against 112 for DD-LS. The benefit is, however, narrow. DD-LS is faster at every size, and by a wide margin. It also leaves smaller gaps, 0.27% against 1.01%, 1.20% and 2.91% on the instances with 20 demand points, where it additionally closes more instances than any of the three versions. V3, which imposes the largest number of linearization constraints upfront, is the weakest of the three throughout. A practitioner with a fixed time budget would therefore choose the original cut, and would turn to the combination only when the objective is to close the gap on instances of moderate size that the original cut leaves open.

An explanation for the gap left by V3 on the largest instances, and for the poor performance of V1 and V2, is to be found in the size of (RMP) once the linearization variables and constraints are added. That size is reported in Appendix A. The central issue is the number of bilinear terms $y_{ij_1}u_{j_2z}$, each of which is replaced by a variable $\ell_{ij_1j_2z}$. Their number is grows quadratically in the number of demand points. It affects all three versions since each of them adds the same ℓ

Table 3: Quality of the bounds. Each row is a separate pairwise comparison between DD-LS and one version, restricted to the `n` instances that neither of the two solved to optimality. Hence, the three rows refer to three different sets of instances. In each pair of columns, DD-LS is the algorithm with (OC) and V* is the version named in the row, that is V1, V2 or V3. `nCutRounds` is the number of cut separation rounds, each adding one cut for DD-LS and two for V* (i.e., cut (28) together with (OC)). `rootDist` and `endDist` are the relative distances of the bound at the end of the root node and at termination, respectively, from the best solution found by any of the methods. That is, both are averages, over the `n` instances of the row, of $100(z^* - L)/|z^*|$, where z^* is the value of the best solution found on that instance by any of the methods and L is the bound in question. Note that z^* is common to the two methods being compared, so that only their bounds differ. Furthermore, z^* is not the incumbent of either of them.

		`nCutRounds`		`rootDist`[%]		`endDist`[%]	
	`n`	DD-LS	V*	DD-LS	V*	DD-LS	V*
V1	53	1581	555	4.64	4.83	0.31	1.11
V2	49	1698	554	4.73	4.75	0.34	1.23
V3	62	1360	187	4.34	4.41	0.27	2.04

variables. With $|\mathcal{J}| = 20$ and $Z = 5$ (RMP) carries on average 17068 variables against the 268 of DD-LS. V1 imposes two inequalities per product and reaches 33881 constraints on average, while V2 and V3 impose one and reach 17082 and 17081 respectively, against 281 for DD-LS. What distinguishes the three versions in practice is therefore not the size of (RMP), which is comparable, but which of the Fortet inequalities is imposed and which is separated, and the consequences this has for how quickly the search can make progress.

To better explain the results it is useful to assess the difference in terms of dual bounds, which we report in Table 3. The comparison made there is pairwise in the sense that each row of the table compares DD-LS with one single version, and is therefore restricted to the instances that neither of those two closes to optimality. On those instances the bound at the end of the root node is essentially the same for the two methods being compared. However, the bound at termination is significantly better for DD-LS. The reason is visible in the `nCutRounds` columns. Each round of separation is more expensive once the bilinear terms are present and the linear programming relaxations involved are larger due to the linearization variables and constraints. It follows that, within the same time limit, DD-LS performs roughly three times as many separation rounds as V1 and V2, and more than seven times as many as V3. An individual combined cut applied by V1-V3 is tighter, in the sense made precise in Proposition 3, but the algorithm is able to generate significantly fewer of them. It follows that the reconstruction of the expected recourse function proceeds more slowly.

Finally, the results illustrate that separating (OC) alongside (28), rather than the latter alone, is preferable. When we drop (OC), V2 solves 43 fewer instances, and more than triples its average solution time. V3 drops from 87 to 59 and V1 from 114 to 105 solved instances. Since (OC) involves none of the linearization variables, it is inexpensive to separate, and it

bounds μ on exactly those solutions for which the bilinear cut is weak.

5 Conclusions

In this article, we introduce a model that considers decision-dependent uncertain demand in single-source capacitated facility location problems. The model is built on the idea of having the distribution of demand dependent on the facility to which demand points are allocated. By partitioning the set of facilities, the model can capture, at one extreme, the case where demands are not influenced by allocation decisions, and, at the other extreme, the case where each individual allocation decision determines a different distribution. Between these extremes, one can find cases where the distribution depends on the subset of facilities to which the demand point is allocated.

To address the exponential size of the model and the possible number of probability distributions, we propose a cutting-plane algorithm based on the L-Shaped method for problems with decision-dependent uncertainty (see Pantuso and Hewitt [2025]), which generates distribution-specific cuts. In addition, we derive a new, tighter, optimality cut based on a bilinear reformulation of the original cut as well as a tight valid inequality.

Extensive numerical tests demonstrate that the method, supported by the valid inequality, can find optimal or near optimal solutions to problem instances with up to 6^{20} potential distributions. This improves significantly the current ability to solve location problems with decision-dependent uncertainty. We also assess the practical value of the new optimality cut. Separated alongside the original one it closes slightly more instances. However, the original cut remains preferable overall, both in terms of solution time and optimality gap attained. The main reason is that the bilinear terms make each round of separation considerably more expensive and the algorithm therefore affords substantially fewer cuts within the same time limit.

6 Acknowledgments

This study was partially funded by the Novo Nordisk Foundation grant NNF24OC0089770.

References

R.K. Ahuja, J.B. Orlin, S. Pallottino, M.P. Scaparra, and M.G. Scutellà. A multi-exchange heuristic for the single-source capacitated facility location problem. *Management Science*, 50 (6):749–760, 2004.

R.M. Apap and I.E. Grossmann. Models and computational strategies for multistage stochastic programming under endogenous and exogenous uncertainties. *Computers & Chemical Engineering*, 103:233–274, 2017.

D. Arthur and S. Vassilvitskii. k-means++: The advantages of careful seeding. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '07, page 1027–1035, USA, 2007. Society for Industrial and Applied Mathematics.

B. Basciftci, S. Ahmed, and S. Shen. Distributionally robust facility location problem under decision-dependent stochastic demand. *European Journal of Operational Research*, 292(2): 548–561, 2021.

I. Correia and F. Saldanha-da Gama. Facility location under uncertainty. *Location Science*, pages 185–213, 2019.

L.F. Escudero, M.A. Garín, J.F. Monge, and A. Unzueta. On preparedness resource allocation planning for natural disaster relief under endogenous uncertainty with time-consistent risk-averse management. *Computers & Operations Research*, 98:84–102, 2018.

Marshall L Fisher, Ramchandran Jaikumar, and Luk N Van Wassenhove. A multiplier adjustment method for the generalized assignment problem. *Management science*, 32(9):1095–1103, 1986.

R. Fortet. Applications de l'algèbre de Boole en recherche opérationnelle. *Revue Française de Recherche Opérationnelle*, 4(14):17–26, 1960.

Gianfranco Guastaroba and Maria Grazia Speranza. A heuristic for bilp problems: the single source capacitated facility location problem. *European Journal of Operational Research*, 238 (2):438–450, 2014.

L. Hellemo, P.I. Barton, and A. Tomasgard. Decision-dependent probabilities in stochastic programs with recourse. *Computational Management Science*, 15(3):369–395, 2018.

K.S. Hindi and K. Pieńkosz. Efficient solution of large scale, single-source, capacitated plant location problems. *Journal of the Operational Research Society*, 50(3):268–274, 1999.

F. Hooshmand and S.A. MirHassani. Efficient constraint reduction in multistage stochastic programming problems with endogenous uncertainty. *Optimization Methods and Software*, 31:359–376, 2016.

F. Hooshmand and S.A. MirHassani. Reduction of nonanticipativity constraints in multistage stochastic programming problems with endogenous and exogenous uncertainty. *Mathematical Methods of Operations Research*, 87(1):1–18, 2018.

T.W. Jonsbråten, R.J.-B. Wets, and D.L. Woodruff. A class of stochastic programs with decision dependent random elements. *Annals of Operations Research*, 82(0):83–106, 1998.

A. Klose and A. Drexl. Facility location models for distribution system design. *European Journal of Operational Research*, 162(1):4–29, 2005.

G. Laporte, F.V. Louveaux, and L. van Hamme. Exact Solution to a Location Problem with Stochastic Demands. *Transportation Science*, 28(2):95–103, 1994.

G. Laporte, S. Nickel, and F. Saldanha-da Gama. *Introduction to Location Science*. Springer, 2019.

M.A. Lejeune and W. Ma. Drone-delivery network for opioid overdose: Nonlinear integer queueing-optimization models and methods. *Operations Research*, 73:86–108, 2025.

M.A. Lejeune, P. Dehghanian, and W. Ma. Profit-based unit commitment models with price-responsive decision-dependent uncertainty. *European Journal of Operational Research*, 314: 1052–1064, 2024.

M.A. Lejeune, F. Margot, and A.D. de Oliveira. Stochastic optimization model with exogenous and decision-dependent uncertainty for medical evacuation. *INFORMS Journal on Computing*, Accepted, 2026.

C.K.Y. Lin. Stochastic single-source capacitated facility location model with service level requirements. *International Journal of Production Economics*, 117(2):439–451, 2009.

T. Liu, F. Saldanha-da Gama, S. Wang, and Y. Mao. Robust stochastic facility location: Sensitivity analysis and exact solution. *INFORMS Journal on Computing*, 34(5):2776–2803, 2022.

F.V. Louveaux. Discrete stochastic location models. *Annals of Operations Research*, 6:21–34, 1986.

G.P. McCormick. Computability of global solutions to factorable nonconvex programs: Part i—convex underestimating problems. *Mathematical Programming*, 10(1):147–175, 1976.

G. Pantuso. A node formulation for multistage stochastic programs with endogenous uncertainty. *Computational Management Science*, 18(3):325–354, 2021.

G Pantuso. Solution of stochastic facility location problems with combinatorially many decision-dependent distributions. *Journal of Combinatorial Optimization*, 50(42), 2025. doi: https://doi.org/10.1007/s10878-025-01371-7.

G. Pantuso and M. Hewitt. The L-shaped method for stochastic programs with decision-dependent uncertainty. *Mathematical Programming*, Accepted, 2025.

S. Peeta, Sibel S.F., D. Gunnec, and K. Viswanath. Pre-disaster investment decisions for strengthening a highway network. *Computers & Operations Research*, 37(10):1708–1719, 2010.

F. Saldanha-da Gama and S. Wang. *Stochastic Facility Location*, pages 123–179. Springer International Publishing, Cham, 2024.

L.V. Snyder. Facility location under uncertainty: A review. *IIE Transactions*, 38(7):547–564, 2006.

K. Tong, Y. Feng, and G. Rong. Planning under Demand and Yield Uncertainties in an Oil Supply Chain. *Industrial & Engineering Chemistry Research*, 51(2):814–834, 2012.

K. Viswanath, S. Peeta, and S.F. Salman. Investing in the links of a stochastic network to minimize expected shortest path. length. Technical report, Purdue University, Department of Economics, 2004.

Z. Yang, F. Chu, and H. Chen. A cut-and-solve based algorithm for the single-source capacitated facility location problem. *European Journal of Operational Research*, 221(3):521–532, 2012.

A Size of the Relaxed Master Problem

B Solution of $Q(y, \xi)$

The primal and dual solutions to $Q(y, \xi)$ can be efficiently obtained using a direct method devised by Laporte et al. [1994]. The method is based on a specialized sorting procedure and is described in Algorithm 2.

input : A solution (x^v, y^v) to (RMP) and the corresponding distribution d^v
output: Primal solution w^v_{ijs} and dual multipliers $\rho^{sd^v v}_{ij}$ and $\gamma^{sd^v v}_i$

for $s \in \mathcal{S}_{d^v}$ **do**
  for $i \in \mathcal{I} : x^v_i = 1$ **do**
    Rank demand points $j \in \mathcal{J}$ in non-increasing order of the values R_{ij};
    Assign demands $\xi_{jd^v s}$ to i (populating variables w^v_{ijs}) in the same order as the values R_{ij} as long as $y^v_{ij} = 1$, $R_{ij} > 0$ and the residual capacity $M_i - \sum_{j \in \mathcal{J}} w^v_{ijs}$ is not exceeded;
    Let λ_{is} be the index $j \in \mathcal{J}$ such that $y^v_{ij} = 1$, $R_{ij} > 0$ and the demand of j can be satisfied only partially by i. If no such index exists, let $\lambda_{is} = |\mathcal{J}| + 1$ with $R_{i,\lambda_{is}} = 0$;
    Let $\rho^{sd^v v}_{ij} = R_{ij} - R_{i,\lambda_{is}}$ if $j < \lambda_{is}$ (on the ranked list), 0 otherwise;
    Let $\gamma^{sd^v v}_i = R_{i,\lambda_{is}}$;
  end
end

Algorithm 2: Procedure to find primal and dual solutions to $Q(y, \xi)$.

Table 4: Size of Relaxed Master Problem for DD-LS and Versions V1-V3 of Algorithm with New Optimality Cuts (28), averaged over the ten instances of each size. The figures refer to the model as built, before any cut is separated, and are therefore the same whether or not (OC) is added alongside (28). Note that V2 now posts (27c) upfront, for the reason given in Section 3.2, and so no longer starts from a master problem of the size of the one of DD-LS.

			#_variables		#_binary		#_constraints	
	$\|\mathcal{Z}\|$	$\|\mathcal{J}\|$	avg	max	avg	max	avg	max
DD-LS	3	10	118.00	129	117.00	128	121.00	131
		15	173.00	189	172.00	188	181.00	196
		20	228.00	249	227.00	248	241.00	261
	4	10	128.00	139	127.00	138	131.00	141
		15	188.00	204	187.00	203	196.00	211
		20	248.00	269	247.00	268	261.00	281
	5	10	138.00	149	137.00	148	141.00	151
		15	203.00	219	202.00	218	211.00	226
		20	268.00	289	267.00	288	281.00	301
V1	3	10	2918.00	3329	117.00	128	5721.00	6531
		15	6473.00	7389	172.00	188	12781.00	14596
		20	11428.00	13049	227.00	248	22641.00	25861
	4	10	3628.00	4139	127.00	138	7131.00	8141
		15	8063.00	9204	187.00	203	15946.00	18211
		20	14248.00	16269	247.00	268	28261.00	32281
	5	10	4338.00	4949	137.00	148	8541.00	9751
		15	9653.00	11019	202.00	218	19111.00	21826
		20	17068.00	19489	267.00	288	33881.00	38701
V2	3	10	2918.00	3329	2917.00	3328	2922.00	3332
		15	6473.00	7389	6472.00	7388	6482.00	7397
		20	11428.00	13049	11427.00	13048	11442.00	13062
	4	10	3628.00	4139	3627.00	4138	3632.00	4142
		15	8063.00	9204	8062.00	9203	8072.00	9212
		20	14248.00	16269	14247.00	16268	14262.00	16282
	5	10	4338.00	4949	4337.00	4948	4342.00	4952
		15	9653.00	11019	9652.00	11018	9662.00	11027
		20	17068.00	19489	17067.00	19488	17082.00	19502
V3	3	10	2918.00	3329	2917.00	3328	2921.00	3331
		15	6473.00	7389	6472.00	7388	6481.00	7396
		20	11428.00	13049	11427.00	13048	11441.00	13061
	4	10	3628.00	4139	3627.00	4138	3631.00	4141
		15	8063.00	9204	8062.00	9203	8071.00	9211
		20	14248.00	16269	14247.00	16268	14261.00	16281
	5	10	4338.00	4949	4337.00	4948	4341.00	4951
		15	9653.00	11019	9652.00	11018	9661.00	11026
		20	17068.00	19489	17067.00	19488	17081.00	19501