The Disaggregated Integer L-shaped Method for the Stochastic Vehicle Routing Problem

Lucas Parada
Robin Legault
Jean-François Côté
Michel Gendreau

January 2023

The Disaggregated Integer L-shaped Method for the Stochastic Vehicle Routing Problem

Lucas Parada[1,2], Robin Legault[1,3], Jean-François Côté[1,4,*], Michel Gendreau[1,2]

[1.] Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT)
[2.] Department of Mathematical and Industrial Engineering, Polytechnique Montréa
[3.] Department of Computer Science and Operations Research, Université de Montréal
[4.] Department of Operations and Decision Systems, Université Laval, Québec, Canada

Abstract. This paper proposes a new integer L-shaped method for solving a family of two-stage stochastic integer programs whose first-stage solutions can decompose into disjoint components. This method uses a new type of optimality cuts and lower bounding functionals that require the recourse function to be monotonic. In a minimization problem, this property stipulates that the recourse cost of a component must always be higher or equal to that of any of its subcomponents. The stochastic vehicle routing problem is particularly well suited to be solved by this new method, as its solutions can be decomposed into a set of routes. We consider the variant of the vehicle routing problem with stochastic demands in which the recourse policy consists of performing a return trip to the depot whenever a vehicle does not have sufficient capacity to accommodate a newly realized customer demand. This work shows that, in general, this policy leads to a non-monotonic recourse function. However, we show that the monotonicity holds when Poisson distributions and other usual distributions model the customer demands. We present several new lower bounds on the recourse that strengthen the lower bounding functionals and significantly speed up the resolution process. Computational experiments on instances from the literature show that the new approach achieves state-of-the-art results.

Keywords: Stochastic programming, integer L-shaped method, monotonic recourse, vehicle routing problem.

_____

* Corresponding author: jean-francois.cote@fsa.ulaval.ca

1 Introduction

In two-stage stochastic programs, first-stage decisions are taken before the realization of uncertain parameters. Once these parameters are revealed, second-stage decisions are taken to correct the first-stage solution. The objective is to find a first-stage solution that minimizes the cost of the first stage and the expected cost of the second stage. These problems have typically been solved using Benders decomposition (Benders, 1962), or equivalently the L-shaped method (Van Slyke and Wets, 1969) when they have continuous variables and a large number of scenarios. For solving programs with binary variables in both stages, researchers historically relied on the integer L-shaped method of Laporte and Louveaux (1993). These methods can be regarded as branch-and-cut (B&C) algorithms where the expectation function of the objective is relaxed and replaced by a variable that is linearly bounded with inequalities. Nowadays, the L-shaped and integer L-shaped methods are considered to be computationally impractical because of their slow convergence. Several techniques have been proposed in the literature to mitigate this problem. A common approach is to add lower bounding functionals (LBFs) (Birge and Wets, 1986) to the model. These inequalities bound the recourse for a much broader range of solutions than optimality cuts. However, many of the most impactful ones are problem-specific, which limits their range of application. Other methods like the logic-based Benders decomposition of Hooker and Ottosson (2003) try to overcome the problem by providing cuts that include only a small subset of variables. These cuts are then effective for a large number of solutions. Other enhancements are reported in Rahmaniani et al. (2017).

This paper introduces a new integer L-shaped method, named the disaggregated integer L-shaped method (DL-shaped method). This method is tailored for a specific class of two-stage stochastic programs in which, given a first-stage solution, the recourse cost can be expressed as a sum of independent recourse functions involving disjoint sets of first-stage variables. The DL-shaped method also requires each of the resulting recourse functions to be monotonic. In the case of a minimization problem, this property means that adding first-stage variables to a component of the recourse function cannot decrease its value. Many two-stage stochastic programs have such structure and property. For example, in the bin packing problem with uncertain weights, it is possible to decompose the first-stage solution by bin. This problem is found in the allocation of surgeries to operating rooms (Denton et al., 2010). Similarly, in job scheduling problems (Li et al., 2022; Elçi and Hooker, 2022), jobs are assigned to machines, and their first-stage solution can be decomposed by machine. In the latter application, the monotonicity of the objective function corresponds to the fact that assigning more jobs to a machine cannot decrease the expected makespan of its operations.

We propose an implementation of the DL-shaped method for the vehicle routing problem with stochastic demands (VRPSD). In this problem, each customer has to be visited exactly once by a fleet of capacitated vehicles. This structure makes it possible to decompose the recourse cost by route. Customer demands are modeled by independent random variables (RVs) with known distributions and are only observed when a vehicle arrives at the customer's location. A failure occurs when a vehicle arrives at a customer's location with a remaining capacity smaller than the customer's demand. When this happens, a recourse action must be implemented to satisfy the current and next customer demands. This work applies the detour-to-depot (DTD) recourse action, which performs a back-and-forth trip from the customer to the depot to restock the vehicle. We will show that this recourse action has the property of monotonicity under some conditions.

The contributions of this paper are as follows. First, we introduce the DL-shaped method in the context of the VRPSD. Second, we demonstrate that, although the DTD recourse function is not monotonic in general, it is for Poisson distributions and some normal, binomial, Erlang, and negative

binomial distributions when the sum of the expected demands on each route respects the vehicle capacity. Third, we present several new lower bounds that take advantage of the monotonicity of the recourse and are used by the LBFs. Finally, numerical results show that each of the new lower bounds greatly improves over the previous ones found in the literature and that our new method achieves state-of-the-art results on existing benchmark instances. This leads us to propose new instances that are more challenging than the existing ones from the literature.

The remainder of this paper is organized as follows. Section 2 presents the literature on stochastic vehicle routing problems. Section 3 presents the mathematical formulation of the problem. Section 4 presents theoretical results on the monotonicity of the recourse function. Section 5 details the implementation of the new method as well as the new lower bounds for the problem. Section 7 presents the numerical results for our implementation, and lastly, Section 8 presents the conclusion.

2   Related literature

The stochastic vehicle routing problem with random customer demands and multiple depots by Tillman (1969) is considered to be the first studied version of the VRPSD. Nowadays, the state-of-the-art exact methods can solve instances with more than 100 customers (Florio et al., 2020). By this measure, since Tillman (1969), the field has linearly increased over time in its solution capability. Although this is considerable progress for an NP-Hard problem like the VRPSD, Gendreau et al. (2016) conclude that we are still in the early stages of the field, with room for significant improvement in exact solution methods. In this section, we present a historical review of the exact methods developed for the VRPSD.

The a priori and re-optimization paradigms are the two modeling paradigms used in the literature. The choice of either of these paradigms depends on when the stochastic parameters are revealed in the solution process. Formalized by Bertsimas et al. (1990), the a priori paradigm relies on the design of routes that cannot be modified after they have been fixed. Re-optimization was formalized by Dror et al. (1989) as a Markov decision process (MDP) and allows for the construction and adjustment of routes as information is gradually revealed. In recent decades, the increasing availability of information technologies has simplified the gathering of real-time information that helps decision-makers re-optimize a priori routes on the fly. As a consequence, new paradigms have emerged, such as the dynamic and online vehicle routing problems, where the parameter are updated throughout the solution process. A comprehensive survey on the dynamic vehicle routing problem is presented in Pillac et al. (2013).

The DTD and optimal restocking (OR) policies are the most common in the literature. Under the DTD policy, restocking trips are only allowed in case of failure. Studies of the VRPSD with this policy include the works of Gendreau et al. (1995), Laporte et al. (2002), and Jabali et al. (2014). For the OR policy, vehicles are allowed to perform preventive returns to the depot to avoid costly failures. Yang et al. (2000) have shown that the OR policy can be calculated by solving a Bellman equation. Relevant works under this policy include Louveaux and Salazar-González (2018), Salavati-Khoshghalb et al. (2019), Florio et al. (2020), and Hoogendoorn and Spliet (2022). Additionally, variants of these two policies have also been studied. For example, Ak and Erera (2007) propose a recourse policy that is based on pairwise sets of a priori routes. Also, in Secomandi and Margot (2009), an a priori planned tour is partly re-optimized using a MDP that is limited to a restricted set of states. Florio et al. (2022) apply another type of partial re-optimization in which the order of visit of successive customers in a planned route is allowed to be switched.

Numerical experiments by Florio et al. (2020) indicate that the OR policy can lead to savings of up to 5.2% compared to the DTD policy when the capacity constraints are relaxed and the number of vehicles is fixed. In this case, the partial re-optimization policy of Florio et al. (2022) further reduces the expected costs by 0.73% on average. However, Florio et al. (2020) found that the benefit of adopting the OR policy in place of the DTD policy is negligible when capacity constraints are imposed and the number of vehicles is not fixed.

Most exact methods from the literature for solving the VRPSD use the integer L-shaped method. Its main advantages are that it can efficiently handle solutions that include long routes, and it is compatible with several demand distributions. In practice, it is unfortunately limited to instances with loose capacity constraints and requiring at most two to three vehicles. Branch-and-price (B&P) methods have shown to be more efficient for instances with tight capacity constraints and requiring a large number of vehicles. However, they have been almost exclusively implemented for demands following Poisson distributions. Additionally, B&P methods struggle to handle instances with long routes. To overcome the limitations of the integer L-shaped method, tight cuts have been deemed essential to bound the recourse function (Gendreau et al., 2014). The literature has therefore devoted many efforts to propose efficient LBFs (Laporte et al., 2002; Jabali et al., 2014; Hoogendoorn and Spliet, 2022). To our knowledge, no numerical studies have been done to compare the tightness and effectiveness of different LBFs for a given recourse policy.

The idea of disaggregating the recourse, i.e., separating it into independent disjoint functions, is not new, as it was introduced by Wets (1966). In the context of the VRPSD, Séguin (1996) proposed to disaggregate the recourse in terms of routes. Linearly bounding disjoint portions of the recourse cost proved to be efficient for a variant of the VRPSD in Côté et al. (2020). Their numerical results showed that these disjoint bounds help at converging to the optimal solution in fewer iterations. The DL-shaped method builds on these previous works by exploiting new optimality cuts and LBFs that simultaneously bound the value of multiple components of the recourse and apply to a larger number of first-stage solutions.

As noted in the introduction, the DL-shaped method requires the monotonicity of the recourse function. The topic of monotonicity received little attention in the VRPSD literature. To the best of our knowledge, the articles by Dror and Ball (1987) and Kreimer and Dror (1990) constitute the only theoretical works on the matter. They both give sufficient conditions for the probability of a failure to increase at each customer on a route. Although we show that this property is sufficient for the recourse function to be monotonic under the DTD policy, the results of Dror and Ball (1987) and Kreimer and Dror (1990) only apply to independent and identically distributed (i.i.d.) demands. This paper introduces sufficient conditions for the monotonicity of the recourse function that can be applied in the independent but not identically distributed case. We show conditions under which it is respected for Poisson, normal, binomial, Erlang, and negative binomial demands.

3   Mathematical formulation

This section introduces the notation that will be used throughout the paper and the mathematical model of the problem.

The VRPSD is defined as follows. Let $G = \{N_0, E\}$ be an undirected graph where $N_0 = \{0, 1, ..., n\}$ is the set of nodes, with node 0 being the depot, $N = \{1, ..., n\}$ the set of customer nodes, and $E = \{(i, j) : i, j \in N, i < j\}$ the set of edges. Traveling on edge (i, j) incurs a travel cost of c_{ij}. Each customer $i \in N$ has a non-negative demand given by the RV ξ_i, with an expected value μ_i and a standard deviation σ_i. It is assumed that the demand variables are independently

distributed. A fleet of identical vehicles is available to satisfy customer requests. Each vehicle has a capacity Q and must follow a route that starts and ends at the depot. Each customer must be visited exactly once. A route is defined as a sequence of customers starting and ending at the depot, and a sequence is feasible if the sum of the expected demands respects the vehicle capacity. This condition was added by Laporte et al. (2002) to avoid routes that systematically fail while having others that are underutilized. Let $M \subseteq \mathbb{N}$ be the set specifying the number of vehicles that can be used in a solution. For unlimited fleet problems, this set is given by $M = \{\lceil (\sum_{i \in S} \mu_i)/Q \rceil, \ldots, |N|\}$, and when exactly $\bar{m}$ vehicles must be used, $M = \{\bar{m}\}$.

Let $E(S)$ be the edges with both endpoints in S and $E(h)$ the set of edges linked with node h. The VRPSD can be formulated as follows. The variable x_{ij} indicates the number of times edge $(i, j) \in E$ is traversed. The variables associated with pairs of customers are binary, and those linking the depot to a customer can take the values 0, 1, or 2. The variable x_{0i} equals 2 if a vehicle serves a single customer $i \in E$. A binary variable z_m decides whether m vehicles are used in the solution. Let $\mathcal{Q}(x)$ denote the expected recourse cost of the first-stage solution $x = (x_{ij})$. The model is:

$$\min \sum_{(i,j) \in E} c_{ij} x_{ij} + \mathcal{Q}(x) \tag{1}$$

$$\text{s.t.} \sum_{i \in N} x_{0i} = \sum_{m \in M} 2m z_m, \tag{2}$$

$$\sum_{(i,j) \in E(h)} x_{ij} = 2 \qquad\qquad h \in N, \tag{3}$$

$$\sum_{(i,j) \in E(S)} x_{ij} \leq |S| - \left\lceil \frac{\sum_{i \in S} \mu_i}{Q} \right\rceil \qquad\qquad S \subseteq N, \tag{4}$$

$$\sum_{m \in M} z_m = 1, \tag{5}$$

$$x_{ij} \in \{0, 1\} \qquad\qquad (i, j) \in E(N), \tag{6}$$
$$x_{ij} \in \{0, 1, 2\} \qquad\qquad (i, j) \in E(0). \tag{7}$$

The objective (1) is to minimize the sum of travel costs plus the expected recourse cost. Constraint (2) imposes that m routes must be connected to the depot if m vehicles are used. Constraints (3) ensure that customers are visited exactly once. Constraints (4) impose that the expected demand on each route does not exceed the vehicle capacity and that the routes are connected to the depot. Constraint (5) ensures that the number of vehicles that are used in the solution is an element of M. Constraints (6) and (7) define the domain of the variables.

For the DTD recourse policy, it is possible to separate the calculation of the recourse by route. Let $\mathcal{R}^\nu$ be the set of routes in solution x^ν. Then, the expected recourse cost is defined as:

$$\mathcal{Q}(x^\nu) = \sum_{r \in \mathcal{R}^\nu} \mathcal{Q}(r), \tag{8}$$

where $\mathcal{Q}(r) = \min\{\mathcal{Q}^1(r), \mathcal{Q}^2(r)\}$ is the minimum expected recourse cost of both orientations of route $r = (0, i_1, \ldots, i_t, 0)$. The computation of both orientations is required because model (1)-(7) does not capture routes orientation. The expected recourse cost is denoted by $\mathcal{Q}^1(r)$ and $\mathcal{Q}^2(r)$ in the two orientations. Under the DTD strategy, a cost of $2c_{0i}$ is incurred when the demand of

customer i cannot be satisfied. The expected recourse of route r, in the first orientation, is thus calculated as follows.

$$\mathcal{Q}^1(r) = 2 \sum_{j=1}^{t} \sum_{l=1}^{\infty} \mathbb{P}\left(\sum_{k=1}^{j-1} \xi_k \leq lQ < \sum_{k=1}^{j} \xi_k\right) c_{0j} \tag{9}$$

Note that a restocking trip occurs only when a customer's demand exceeds the vehicle's residual capacity. We exclude the case of going to the depot and then moving to the next customer when an exact stock-out occurs. If it is assumed that $\xi_i \leq Q$ for all i with probability 1, then the second summation of (9) can be limited to $j-1$ instead of going to ∞. If demands can exceed the capacity of the vehicles, the probability $\mathbb{P}\left(\sum_{k=1}^{j-1} \xi_k \leq lQ \leq \sum_{k=1}^{j} \xi_k\right)$ tends to zero as l increases and the summation can be stopped when it reaches a probability close to zero. In our implementation, the summation is stopped when the probability is lower than 0.0001. Throughout the paper, the expected recourse cost $\mathcal{Q}(r)$ of a route $r = (0, p, 0)$ will equivalently be denoted by $\mathcal{Q}(p)$.

4   Monotonicity of the recourse function

Our algorithm relies on new optimality cuts and lower bounds that require the recourse function to decrease monotonically when removing customers from a route. This section formalizes this property by the so-called monotonicity condition and monotonicity property, which respectively apply to sets of customers and instances of the VRPSD. Based on these definitions, we prove sufficient conditions for the recourse function to be monotonic under the DTD policy when the demands follow independent Poisson, normal, binomial, Erlang, and negative binomial distributions. We also provide counterexamples to illustrate that the recourse function is not monotonic in general.

Definition 1. *Let $p = (i_1, \ldots, i_t)$ be a path on graph G. We say that $p' = (j_1, \ldots, j_{t'})$ is a subpath of p if it can be obtained by removing elements of p, without changing the ordering of the remaining elements. Formally, p' is thus said to be a subpath of p if $\{j_1, \ldots, j_{t'}\} \subseteq \{i_1, \ldots, i_t\}$ and, for any pair $(j_{a'}, j_{b'}) = (i_a, i_b)$, $a' < b'$ if and only if $a < b$.*

Definition 2. *Let $p = (i_1, \ldots, i_t)$ be a path on graph G. A subpath p' that can be written as $p' = (i_a, \ldots, i_b)$ for some $1 \leq a \leq b \leq t$ is said to be connected. In other words, a connected subpath p' is obtained by removing elements at the beginning and the end of p.*

Definition 3. *A set of customers $S \subseteq N$ is said to respect the monotonicity condition if, for any pair $(a, b) \in S \times S$ such that $a \neq b$ and for any subset $\tilde{S} \subseteq S \setminus \{a, b\}$, the following inequality:*

$$\mathbb{P}\left(\tilde{\xi} + \xi_a \leq lQ < \tilde{\xi} + \xi_a + \xi_b\right) \geq \mathbb{P}\left(\tilde{\xi} \leq lQ < \tilde{\xi} + \xi_b\right), \tag{10}$$

where $\tilde{\xi} = \sum_{i \in \tilde{S}} \xi_i$, holds for any positive integer $l \in \mathbb{N}$.

Definition 4. *An instance of the VRPSD with the DTD policy is said to have the monotonicity property if any set $S \subseteq N$ such that $\sum_{i \in S} \mu_i \leq Q$ respects the monotonicity condition.*

The above definition can be applied by enumeration to verify whether an instance respects the monotonicity property, but it might be unpractical for large instances since the number of comparisons grows exponentially with $|N|$. In Propositions 1 and 2, we introduce an equivalent formulation of the monotonicity condition and a sufficient condition which will later be used to prove that the monotonicity property is respected if the demands are modeled by specific distributions.

Proposition 1. *Inequality (10) can be rewritten equivalently as follows.*

$$\mathbb{P}\left(\tilde{\xi} + \xi_a \leq lQ\right) - \mathbb{P}\left(\tilde{\xi} + \xi_a + \xi_b \leq lQ\right) \geq \mathbb{P}\left(\tilde{\xi} \leq lQ\right) - \mathbb{P}\left(\tilde{\xi} + \xi_b \leq lQ\right) \qquad (11)$$

Proof.

$$\mathbb{P}\left(\tilde{\xi} + \xi_a \leq lQ < \tilde{\xi} + \xi_a + \xi_b\right) \geq \mathbb{P}\left(\tilde{\xi} \leq lQ < \tilde{\xi} + \xi_b\right) \qquad (12)$$

$$\iff \mathbb{P}\left(\tilde{\xi} + \xi_a + \xi_b > lQ\right) - \mathbb{P}\left(\tilde{\xi} + \xi_a > lQ\right) \geq \mathbb{P}\left(\tilde{\xi} + \xi_b > lQ\right) - \mathbb{P}\left(\tilde{\xi} > lQ\right) \qquad (13)$$

$$\iff \mathbb{P}\left(\tilde{\xi} + \xi_a \leq lQ\right) - \mathbb{P}\left(\tilde{\xi} + \xi_a + \xi_b \leq lQ\right) \geq \mathbb{P}\left(\tilde{\xi} \leq lQ\right) - \mathbb{P}\left(\tilde{\xi} + \xi_b \leq lQ\right) \qquad (14)$$

Inequalities (12) and (13) are equivalent since it is assumed that the demands have non-negative support. $\qquad \square$

Proposition 2. *Let $S \subseteq N$ be a set of customers such that, for any pair $(a, b) \in S \times S$, $a \neq b$, and for any subset $\tilde{S} \subseteq S \setminus \{a, b\}$, the random variables $\tilde{\xi}$, ξ_a and ξ_b can respectively be rewritten as the sum of $\tilde{\rho} \in \mathbb{N}$, $\rho_a \in \mathbb{N}$ and $\rho_b \in \mathbb{N}$ i.i.d. non-negative RVs ζ_k, $k \in \mathbb{N}$. The following condition is sufficient for the set S to respect the monotonicity condition:*

$$P_{lQ}(\tilde{\rho} + \rho_a + j) \geq P_{lQ}(\tilde{\rho} + j) \quad \forall l \in \mathbb{N}, \forall j \in \{1, \ldots, \rho_b\}, \qquad (15)$$

where $P_u(s) = \mathbb{P}\left(\sum_{k=1}^{s-1} \zeta_k \leq u < \sum_{k=1}^{s} \zeta_k\right)$ denotes the probability that exactly $s \in \mathbb{N}$ RVs ζ_k must be drawn for their cumulative sum to first exceed a given threshold $u \geq 0$.

Proof. Given the assumptions of the proposition, each original customer $i \in S$ can be seen as a group of ρ_i smaller customers with i.i.d. demands ζ_k, $k \in \mathbb{N}$ and sharing the same location.

From there, the left-hand side of inequality (10) can be rewritten as follows.

$$\mathbb{P}\left(\tilde{\xi} + \xi_a \leq lQ < \tilde{\xi} + \xi_a + \xi_b\right) = \mathbb{P}\left(\sum_{k=1}^{\tilde{\rho}+\rho_a} \zeta_k \leq lQ < \sum_{k=1}^{\tilde{\rho}+\rho_a+\rho_b} \zeta_k\right)$$

$$= \sum_{j=1}^{\rho_b} \mathbb{P}\left(\sum_{k=1}^{\tilde{\rho}+\rho_a+j-1} \zeta_k \leq lQ < \sum_{k=1}^{\tilde{\rho}+\rho_a+j} \zeta_k\right)$$

$$= \sum_{j=1}^{\rho_b} P_{lQ}(\tilde{\rho} + \rho_a + j)$$

The second equality is obtained by decomposing the probability that the l-th restocking trip occurs between customers $\tilde{\rho}+\rho_a+1$ and $\tilde{\rho}+\rho_a+\rho_b$ as the sum of the probabilities that it occurs at customer $\tilde{\rho}+\rho_a+j$, for $j=1, \ldots, \rho_b$.

Doing the same for the right-hand side, we obtain the sufficient condition for inequality (10).

$$\mathbb{P}\left(\tilde{\xi} + \xi_a \leq lQ < \tilde{\xi} + \xi_a + \xi_b\right) \geq \mathbb{P}\left(\tilde{\xi} \leq lQ < \tilde{\xi} + \xi_b\right)$$

$$\iff \sum_{j=1}^{\rho_b} P_{lQ}(\tilde{\rho} + \rho_a + j) \geq \sum_{j=1}^{\rho_b} P_{lQ}(\tilde{\rho} + j)$$

$$\Longleftarrow P_{lQ}(\tilde{\rho} + \rho_a + j) \geq P_{lQ}(\tilde{\rho} + j) \qquad \forall j \in \{1, \ldots, \rho_b\}$$

$$\square$$

We now prove that, if the customers of a path respect the monotonicity condition, removing customers from this path cannot increase its recourse cost. This is shown in Proposition 3.

Proposition 3. *For a path $p = (i_1, \ldots, i_t)$ composed of a set of customers respecting the monotonicity condition, the expected recourse cost decreases monotonically when customers are removed from p, i.e., $\mathcal{Q}(p) \geq \mathcal{Q}(p')$ for any subpath p' of p.*

Proof.

$$
\mathcal{Q}(p) = 2 \sum_{j=1}^{t} \sum_{l=1}^{\infty} \mathbb{P}\left(\sum_{k=1}^{j-1} \xi_{i_k} \leq lQ < \sum_{k=1}^{j} \xi_{i_k} \right) c_{0i_j}
$$

$$
= 2 \sum_{\substack{j=1 \\ i_j \in p'}}^{t} \sum_{l=1}^{\infty} \mathbb{P}\left(\sum_{k=1}^{j-1} \xi_{i_k} \leq lQ < \sum_{k=1}^{j} \xi_{i_k} \right) c_{0i_j} + 2 \sum_{\substack{j=1 \\ i_j \in p \setminus p'}}^{t} \sum_{l=1}^{\infty} \mathbb{P}\left(\sum_{k=1}^{j-1} \xi_{i_k} \leq lQ < \sum_{k=1}^{j} \xi_{i_k} \right) c_{0i_j}
$$

$$
\geq 2 \sum_{\substack{j=1 \\ i_j \in p'}}^{t} \sum_{l=1}^{\infty} \mathbb{P}\left(\sum_{k=1}^{j-1} \xi_{i_k} \leq lQ < \sum_{k=1}^{j} \xi_{i_k} \right) c_{0i_j}
$$

$$
\geq 2 \sum_{\substack{j=1 \\ i_j \in p'}}^{t} \sum_{l=1}^{\infty} \mathbb{P}\left(\sum_{\substack{k=1 \\ i_k \in p'}}^{j-1} \xi_{i_k} \leq lQ < \sum_{\substack{k=1 \\ i_k \in p'}}^{j} \xi_{i_k} \right) c_{0i_j}
$$

$$
= \mathcal{Q}(p')
$$

The second inequality comes from the fact that, for any $l \geq 1$ and for any $j \in \{1, \ldots, t\}$ such that $i_j \in p'$, the following inequality holds.

$$
\mathbb{P}\left(\sum_{k=1}^{j-1} \xi_{i_k} \leq lQ < \sum_{k=1}^{j} \xi_{i_k} \right) \geq \mathbb{P}\left(\sum_{\substack{k=1 \\ i_k \in p'}}^{j-1} \xi_{i_k} \leq lQ < \sum_{\substack{k=1 \\ i_k \in p'}}^{j} \xi_{i_k} \right) \tag{16}
$$

To prove this, we first introduce some notation:

- $\tilde{S}^j = \{i_k \in p' : k \leq j-1\}$ is the set of customers that are visited before customer i_j in path p'

- $\tilde{\xi}^j = \sum_{i \in \tilde{S}^j} \xi_i$ is the total demand of these customers

- $\{a_1, \ldots, a_s\} = \{i_k \in p : k \leq j-1, i_k \notin p'\}$ is the set of customers that are visited before customer i_j in path p, but are not visited in path p'

We can now prove inequality (16).

$$
\mathbb{P}\left(\sum_{\substack{k=1 \\ i_k \in p'}}^{j-1} \xi_{i_k} \leq lQ < \sum_{\substack{k=1 \\ i_k \in p'}}^{j} \xi_{i_k} \right) = \mathbb{P}\left(\tilde{\xi}^j \leq lQ < \tilde{\xi}^j + \xi_{i_j} \right) \tag{17}
$$

$$\leq \mathbb{P}\left(\tilde{\xi}^j + \xi_{a_1} \leq lQ < \tilde{\xi}^j + \xi_{a_1} + \xi_{i_j}\right) \tag{18}$$

$$\vdots \tag{19}$$

$$\leq \mathbb{P}\left(\tilde{\xi}^j + \xi_{a_1} + \cdots + \xi_{a_s} \leq lQ < \tilde{\xi}^j + \xi_{a_1} + \cdots + \xi_{a_s} + \xi_{i_j}\right) \tag{20}$$

$$= \mathbb{P}\left(\sum_{k=1}^{j-1} \xi_{i_k} \leq lQ < \sum_{k=1}^{j} \xi_{i_k}\right) \tag{21}$$

By hypothesis, the set $S = \{i_1, \ldots, i_j\}$ respects the monotonicity condition. This allows us to apply inequality (10) s times for different pairs $(a, b) \in S \times S$ and subsets $\tilde{S} \subseteq S$ to obtain inequalities (18)-(20). At the q-th application of the inequality, $(a, b) = (a_q, i_j)$ and $\tilde{S} = \tilde{S}^j \cup \{a_1, \ldots, a_{q-1}\}$. $\quad\square$

Propositions 4 and 5 show that, under some conditions, the monotonicity property is respected for customer demands following independent Poisson, normal, binomial, Erlang, and negative binomial distributions.

Proposition 4. *Any set S of customers whose demands are given by independent Poisson RVs $\xi_i \sim Poisson(\lambda_i)$ with total expected value $\sum_{i \in S} \lambda_i \leq Q$ respects the monotonicity condition.*

Proof. Consider a pair $(a, b) \in S \times S$ such that $a \neq b$, a subset $\tilde{S} \subseteq S \setminus \{a, b\}$ with total expected demand $\tilde{\lambda} = \sum_{i \in \tilde{S}} \lambda_i$ and a positive integer $l \in \mathbb{N}$. To show that inequality (11) holds, we use the fact that the sum of s independent Poisson RVs with means $\lambda_1, \ldots, \lambda_s$ is Poisson distributed with mean $\sum_{k=1}^{s} \lambda_k$. Also, we use the fact that the cumulative distribution function (CDF) of a Poisson RV with mean λ evaluated at $lQ \in \mathbb{N}$ is given by $H(lQ, \lambda) = \frac{\Gamma(lQ+1, \lambda)}{(lQ)!}$, where $\Gamma(lQ+1, \lambda) = \int_{\lambda}^{\infty} t^{lQ} e^{-t} dt$ is the upper incomplete gamma function.

$$\mathbb{P}\left(\tilde{\xi} + \xi_a \leq lQ\right) - \mathbb{P}\left(\tilde{\xi} + \xi_a + \xi_b \leq lQ\right) \geq \mathbb{P}\left(\tilde{\xi} \leq lQ\right) - \mathbb{P}\left(\tilde{\xi} + \xi_b \leq lQ\right) \tag{22}$$

$$\iff H\left(lQ, \tilde{\lambda}+\lambda_a\right) - H\left(lQ, \tilde{\lambda}+\lambda_a+\lambda_b\right) \geq H\left(lQ, \tilde{\lambda}\right) - H\left(lQ, \tilde{\lambda}+\lambda_b\right) \tag{23}$$

$$\iff \frac{\Gamma\left(lQ+1, \tilde{\lambda}+\lambda_a\right)}{(lQ)} - \frac{\Gamma\left(lQ+1, \tilde{\lambda}+\lambda_a+\lambda_b\right)}{(lQ)} \geq \frac{\Gamma\left(lQ+1, \tilde{\lambda}\right)}{(lQ)} - \frac{\Gamma\left(lQ+1, \tilde{\lambda}+\lambda_b\right)}{(lQ)} \tag{24}$$

$$\iff \Gamma\left(lQ+1, \tilde{\lambda}+\lambda_a\right) - \Gamma\left(lQ+1, \tilde{\lambda}+\lambda_a+\lambda_b\right) \geq \Gamma\left(lQ+1, \tilde{\lambda}\right) - \Gamma\left(lQ+1, \tilde{\lambda}+\lambda_b\right) \tag{25}$$

$$\iff \int_{\tilde{\lambda}+\lambda_a}^{\infty} \lambda^{lQ} e^{-\lambda} d\lambda - \int_{\tilde{\lambda}+\lambda_a+\lambda_b}^{\infty} \lambda^{lQ} e^{-\lambda} d\lambda \geq \int_{\tilde{\lambda}}^{\infty} \lambda^{lQ} e^{-\lambda} d\lambda - \int_{\tilde{\lambda}+\lambda_b}^{\infty} \lambda^{lQ} e^{-\lambda} d\lambda \tag{26}$$

$$\iff \int_{\tilde{\lambda}+\lambda_a}^{\tilde{\lambda}+\lambda_a+\lambda_b} \lambda^{lQ} e^{-\lambda} d\lambda \geq \int_{\tilde{\lambda}}^{\tilde{\lambda}+\lambda_b} \lambda^{lQ} e^{-\lambda} d\lambda \tag{27}$$

$$\iff \int_{\tilde{\lambda}}^{\tilde{\lambda}+\lambda_b} (\lambda+\lambda_a)^{lQ} e^{-(\lambda+\lambda_a)} d\lambda \geq \int_{\tilde{\lambda}}^{\tilde{\lambda}+\lambda_b} \lambda^{lQ} e^{-\lambda} d\lambda \tag{28}$$

$$\iff \int_{\tilde{\lambda}}^{\tilde{\lambda}+\lambda_b} \left((\lambda+\lambda_a)^{lQ} e^{-(\lambda+\lambda_a)} - \lambda^{lQ} e^{-\lambda}\right) d\lambda \geq 0 \tag{29}$$

$$\iff \int_{\tilde{\lambda}}^{\tilde{\lambda}+\lambda_b} \left(f_{lQ}(\lambda+\lambda_a) - f_{lQ}(\lambda)\right) d\lambda \geq 0, \tag{30}$$

8

where $f_{lQ}(\lambda) := \lambda^{lQ}e^{-\lambda}$. Since the inequalities $0 \leq \lambda \leq \lambda + \lambda_a \leq Q \leq lQ$ hold for any $\lambda \in [\tilde{\lambda}, \tilde{\lambda} + \lambda_b]$, it is sufficient to show that the function $f_{lQ}(\cdot)$ is non-decreasing on $[0, lQ]$ to conclude that the difference $f_{lQ}(\lambda + \lambda_a) - f_{lQ}(\lambda)$ is non-negative everywhere on the interval of integration. The derivative of function $f_{lQ}(\cdot)$ is given by:

$$\frac{\partial f_{lQ}(\lambda)}{\partial \lambda} = (lQ - \lambda)e^{-\lambda}\lambda^{lQ-1},$$

which is indeed non-negative $\forall \lambda \in [0, lQ]$. Inequality (30) is therefore respected. $\square$

Proposition 5. *Any set S of customers whose demands are given by independent normal RVs $\xi_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$ with integer expected values $\mu_i \in \mathbb{N}$ and sharing a common coefficient of dispersion $D = \frac{\sigma_i^2}{\mu_i} \leq 1 \ \forall i \in S$ with total expected value $\sum_{i \in S} \mu_i \leq Q$ respects the monotonicity condition.*

Proof. Although the support of a normally distributed RV is $\mathbb{R}$, the demands are regarded as non-negative in the proof. For this approximation not to be problematic in practice, the standard deviation σ_i of each customer $i \in N$ should not exceed a small fraction of its demand μ_i.

Consider a pair $(a, b) \in S \times S$ such that $a \neq b$ and a subset $\tilde{S} \subseteq S \setminus \{a, b\}$ with total expected demand $\tilde{\mu} = \sum_{i \in \tilde{S}} \mu_i$. Using the notation of Proposition 2, $\tilde{\xi}$, ξ_a and ξ_b can respectively be decomposed as the sum of $\tilde{\rho} = \tilde{\mu}$, $\rho_a = \mu_a$ and $\rho_b = \mu_b$ i.i.d. RVs ξ_k, $k \in \mathbb{N}$, such that $\zeta_k \sim \mathcal{N}(\mu = 1, \sigma^2 = D)$. From there, the sufficient condition (15) can be rewritten as follows.

$$P_{lQ}(\tilde{\mu} + \mu_a + j) \geq P_{lQ}(\tilde{\mu} + j) \quad \forall l \in \mathbb{N}, \forall j \in \{1, \dots, \mu_b\} \tag{31}$$

The coefficient of variation $c = \sigma/\mu$ of the i.i.d. RVs ζ_k is equal to $\sqrt{D}$. Since $D \leq 1$ by hypothesis, then $c \leq 1$. In this case, it follows from Kreimer and Dror (1990) that, for any $l \in \mathbb{N}$, the sequence $\{P_{lQ}(s)\}_{s=1}^{\lfloor lQ \rfloor}$ is monotonically increasing. Since $\tilde{\mu} + \mu_a + \mu_b \leq \sum_{i \in S} \mu_i \leq Q$ and all the expected demands are integer-valued, this allows to conclude that (31) holds. $\square$

Proposition 6. *Any set S of customers whose demands are given by independent binomial RVs $\xi_i \sim Bin(n_i, p)$ sharing a common success probability p with total expected value $\sum_{i \in S} n_i p \leq Q$ respects the monotonicity condition.*

Proof. Consider a pair $(a, b) \in S \times S$ such that $a \neq b$ and a subset $\tilde{S} \subseteq S \setminus \{a, b\}$ with total expected demand $\tilde{n}p$, where $\tilde{n} = \sum_{i \in \tilde{S}} n_i$. Using the notation of Proposition 2, the demands $\tilde{\xi}$, ξ_a and ξ_b can respectively be decomposed as the sum of $\tilde{\rho} = \tilde{n}$, $\rho_a = n_a$ and $\rho_b = n_b$ i.i.d. RVs ξ_k, $k \in \mathbb{N}$, such that $\zeta_k \sim \text{Bern}(p)$. From there, the sufficient condition (15) can be rewritten as follows.

$$P_{lQ}(\tilde{n} + n_a + j) \geq P_{lQ}(\tilde{n} + j) \quad \forall l \in \mathbb{N}, \forall j \in \{1, \dots, n_b\} \tag{32}$$

It follows from Kreimer and Dror (1990) that, for any $l \in \mathbb{N}$, the sequence $\{P_{lQ}(s)\}_{s=1}^{\lceil \frac{lQ}{p} \rceil}$ is non-decreasing. Since $\tilde{n} + n_a + n_b \leq \sum_{i \in S} n_i \leq Q/p$, this allows to conclude that (32) holds. $\square$

Proposition 7. *Any set S of customers whose demands are given by independent Erlang RVs $\xi_i \sim Erlang(n_i, \lambda)$ sharing a common rate parameter λ with total expected value $\sum_{i \in S} n_i/\lambda \leq Q$ respects the monotonicity condition.*

Proof. Consider a pair $(a, b) \in S \times S$ such that $a \neq b$ and a subset $\tilde{S} \subseteq S \setminus \{a, b\}$ with total expected demand $\tilde{n}/\lambda$, where $\tilde{n} = \sum_{i \in \tilde{S}} n_i$. Using the notation of Proposition 2, the demands $\tilde{\xi}$, ξ_a and ξ_b can respectively be decomposed as the sum of $\tilde{\rho} = \tilde{n}$, $\rho_a = n_a$ and $\rho_b = n_b$ i.i.d. RVs ξ_k, $k \in \mathbb{N}$, such that $\zeta_k \sim \text{Exp}(\lambda)$. From there, the sufficient condition (15) can be rewritten as follows.

$$P_{lQ}(\tilde{n} + n_a + j) \geq P_{lQ}(\tilde{n} + j) \quad \forall l \in \mathbb{N}, \forall j \in \{1, \ldots, n_b\} \tag{33}$$

It follows from Kreimer and Dror (1990) that, for any $l \in \mathbb{N}$, the sequence $\{P_{lQ}(s)\}_{s=1}^{\lceil \lambda l Q \rceil}$ is non-decreasing. Since $\tilde{n} + n_a + n_b \leq \sum_{i \in S} n_i \leq \lambda Q$, this allows to conclude that (33) holds. $\quad\square$

Proposition 8. *Any set S of customers whose demands are given by independent negative binomial RVs $\xi_i \sim NB(r_i, p)$ sharing a common success probability p with total expected value $\sum_{i \in S} r_i(1-p)/p \leq Q$ respects the monotonicity condition.*

Proof. Consider a pair $(a, b) \in S \times S$ such that $a \neq b$, a subset $\tilde{S} \subseteq S \setminus \{a, b\}$ with total expected demand $\tilde{r}(1-p)/p$, where $\tilde{r} = \sum_{i \in \tilde{S}} r_i$ and a positive integer $l \in \mathbb{N}$. We use the fact that the sum of s independent negative binomial RVs with parameters $r_1, \ldots, r_s$ and the same success probability p is also a negative binomial RV with parameters $r = \sum_{k=1}^{s} r_k$ and p to rewrite inequality (10) as follows.

$$\mathbb{P}\left(\tilde{\xi} + \xi_a \leq lQ < \tilde{\xi} + \xi_a + \xi_b\right) \geq \mathbb{P}\left(\tilde{\xi} \leq lQ < \tilde{\xi} + \xi_b\right) \tag{34}$$

$$\iff \mathbb{P}\left(\sum_{k=1}^{lQ+\tilde{r}+r_a} \zeta_k \leq lQ < \sum_{k=1}^{lQ+\tilde{r}+r_a+r_b} \zeta_k\right) \geq \mathbb{P}\left(\sum_{k=1}^{lQ+\tilde{r}} \zeta_k \leq lQ < \sum_{k=1}^{lQ+\tilde{r}+r_b} \zeta_k\right) \tag{35}$$

To obtain inequality (35), we use the fact that the CDF of a negative binomial with parameters r and p evaluated at $lQ \in \mathbb{N}$ is equal to the CDF of a binomial distribution with $lQ + r$ trials and probability of success $1 - p$ evaluated at lQ. In inequality (35), these binomial RVs are decomposed into sums of i.i.d. RVs $\zeta_k \sim \text{Bern}(1-p)$. Using this reformulation of the monotonicity condition based on the RVs ζ_k, it suffices to show that $P_{lQ}(lQ + \tilde{r} + r_a + j) \geq P_{lQ}(lQ + \tilde{r} + j) \, \forall j \in \{1, \ldots, r_b\}$ to conclude the proof, by Proposition 2. For a given $j \in \{1, \ldots, r_b\}$, this inequality can be developed as follows:

$$P_{lQ}(lQ + \tilde{r} + r_a + j) \geq P_{lQ}(lQ + \tilde{r} + j) \tag{36}$$

$$\iff \mathbb{P}\left(\sum_{k=1}^{lQ+\tilde{r}+r_a+j-1} \zeta_k \leq lQ < \sum_{k=1}^{lQ+\tilde{r}+r_a+j} \zeta_k\right) \geq \mathbb{P}\left(\sum_{k=1}^{lQ+\tilde{r}+j-1} \zeta_k \leq lQ < \sum_{k=1}^{lQ+\tilde{r}+j} \zeta_k\right) \tag{37}$$

$$\iff \mathbb{P}\left(\chi \leq lQ < \chi + \zeta_1\right) \geq \mathbb{P}\left(\bar{\chi} \leq lQ < \bar{\chi} + \zeta_1\right), \tag{38}$$

where $\chi \sim \text{Bin}(lQ + \tilde{r} + r_a + j - 1, 1 - p)$ and $\bar{\chi} \sim \text{Bin}(lQ + \tilde{r} + j - 1, 1 - p)$. Note that the event in the left-hand side probability of inequality (38) occurs if and only if $\chi = lQ$ and $\zeta_1 = 1$. Since χ and ζ_1 are independent, the joint probability of these two events is equal to the product of their probabilities. Applying the same reasoning to the left-hand side, inequality (38) can be written as follows.

$$\mathbb{P}\left(\chi = lQ\right) \mathbb{P}\left(\zeta_1 = 1\right) \geq \mathbb{P}\left(\bar{\chi} = lQ\right) \mathbb{P}\left(\zeta_1 = 1\right) \tag{39}$$

$$\iff \mathbb{P}\left(\chi = lQ\right) \geq \mathbb{P}\left(\bar{\chi} = lQ\right) \tag{40}$$

$$\iff \frac{(lQ + \tilde{r} + r_a + j - 1)!}{(lQ)!(\tilde{r} + r_a + j - 1)!}(1-p)^{lQ} p^{\tilde{r}+r_a+j-1} \geq \frac{(lQ + \tilde{r} + j - 1)!}{(lQ)!(\tilde{r} + j - 1)!}(1-p)^{lQ} p^{\tilde{r}+j-1} \tag{41}$$

$$\Longleftrightarrow \frac{\prod_{i=1}^{r_a}(lQ + \tilde{r} + i + j - 1)}{\prod_{i=1}^{r_a}(\tilde{r} + i + j - 1)} p^{r_a} \geq 1 \tag{42}$$

$$\Longleftrightarrow \prod_{i=1}^{r_a}\left(\frac{p(lQ + \tilde{r} + i + j - 1)}{\tilde{r} + i + j - 1}\right) \geq 1 \tag{43}$$

To complete the proof, we demonstrate that inequality (43) holds by showing that each term in the product is greater than or equal to 1. For $i \in \{1, \ldots, r_a\}$, we have:

$$\frac{p(lQ + \tilde{r} + i + j - 1)}{\tilde{r} + i + j - 1} = \frac{plQ}{\tilde{r} + i + j - 1} + p \tag{44}$$

$$\geq \frac{pQ}{\tilde{r} + r_a + r_b} + p \tag{45}$$

$$\geq \frac{pQ}{\sum_{i \in S} r_i} + p \tag{46}$$

$$\geq \frac{(1-p)pQ}{pQ} + p \tag{47}$$

$$= 1 \tag{48}$$

To obtain inequality (45), we notice that $l \geq 1$ and $i + j - 1 \leq r_a + r_b - 1 \leq r_a + r_b$. Inequality (47) follows from the assumption that the total expected demand of the customers of set S does not exceed the vehicle capacity. $\square$

The next propositions present two counterexamples to illustrate that the recourse function is not monotonic in general.

Proposition 9. *Let $p = (i_1, \ldots, i_t)$ be a path of customers whose demands are given by independent Poisson RVs $\xi_i \sim Poisson(\lambda_i)$ with total expected value $\sum_{i=1}^{t} \lambda_i > Q$, and p' a subpath of p. Inequality $\mathcal{Q}(p) \geq \mathcal{Q}(p')$ does not hold in general.*

Proof. Consider a problem with $Q = 20$, a path $p = (1, 2, 3)$, and $p' = (2, 3)$. Suppose the expected demand of each customer is given by $\lambda_1 = 5$, $\lambda_2 = 15$, and $\lambda_3 = 10$, respectively. If $c_{01} = c_{02} = 0$ and $c_{03} = 1$, then $\mathcal{Q}(p) \approx 1.11 < 1.47 \approx \mathcal{Q}(p')$. $\square$

Proposition 10. *Let $p = (i_1, \ldots, i_t)$ be a path of customers whose total expected demand respects the vehicle capacity and p' a subpath of p. Inequality $\mathcal{Q}(p) \geq \mathcal{Q}(p')$ does not hold in general.*

Proof. Consider a problem with $Q = 20$, a path $p = (1, 2, 3)$, and $p' = (2, 3)$. Suppose the demand of customer 1 is 5 with probability 1, while that of customers 2 and 3 is 6 with probability 0.9 and 16 with probability 0.1. If $c_{03} > c_{02}$, then $\mathcal{Q}(p) = (0.1)2c_{02} + (0.09)2c_{03} < (0.19)2c_{03} = \mathcal{Q}(p')$. $\square$

5 Disaggregated integer L-shaped method

This section presents the DL-shaped method as a variant of the classical integer L-shaped method from Laporte and Louveaux (1993). First, an outline of the classical method is provided, and we motivate the need for our method by exposing the known shortcoming of the classical one. Next, Section 5.1 presents new optimality cuts, from which the validity of our method results. Finally, Section 5.2 introduces the new LBFs we use to approximate the recourse cost.

The integer L-shaped method for two-stage recourse problems starts by constructing a so-called master problem (MP). The MP is obtained from the original problem by relaxing a set of complicating constraints. This set typically includes the integrality constraints and any set of constraints that is exponential in size, like the classical subtour elimination constraints. In addition, the recourse function is replaced by a continuous variable $\theta \geq 0$ that bounds it from below. If a lower bound on the expected recourse L can be computed, then the inequality $\theta \geq L$ is added to the model. Otherwise, the bound on θ remains the trivial one. The MP is thus a problem defined by the non-relaxed constraints of the original program, the non-negativity constraint on θ, the lower bound L on θ, and an objective function of the form $\min cx + \theta$, where c is a known cost vector with the same size as the solution vector. The problem is then solved with an iterative branch-and-bound method that adds linear inequalities on the fly to explore the set of feasible solutions of the original program. The first MP that is solved in the method is denoted as the root node, and the following MPs are indexed by the iteration at which they appear in the resolution process. The linear inequalities that are added to the successive MPs are either feasibility or optimality cuts. The feasibility cuts prohibit infeasible solutions from being visited, while optimality cuts improve the bound on θ whenever a better feasible solution is found. The feasibility cuts are generally problem-specific. In the context of the VRPSD, they typically correspond to subtour elimination constraints and rounded capacity inequalities. Optimality cuts have a general form regardless of the problem to which the integer L-shaped method is applied. Let x_1^ν be the set of positive variables in a feasible solution x^ν at iteration ν of the method. Then the following inequality:

$$\theta \geq (\mathcal{Q}(x^\nu) - L) \left(\sum_{i \in x_1^\nu} x_i - \sum_{i \notin x_1^\nu} x_i - |x_1^\nu| + 1 \right) + L, \tag{49}$$

where $\mathcal{Q}(x^\nu)$ is the expected recourse cost of solution x^ν, is a valid optimality cut. It was shown by Laporte and Louveaux (1993) that these optimality cuts will yield an optimal solution, if one exists, in a finite number of steps.

The known shortcoming of the integer L-shaped method is that its optimality cuts are active for a single solution. The optimality cut reduces to $\theta \geq \mathcal{Q}(x^\nu)$ at solution x^ν, but for any other solution, it does not improve the inequality $\theta \geq L$. Consequently, without appropriate LBFs to increase the bound on θ, the classical method generally iterates through many first-stage solutions before converging to an optimal solution.

5.1   New optimality cuts

The main idea of the new optimality cuts in the DL-shaped method is to disaggregate the recourse into several components. This idea is not new and can be found, for example, in the work of Birge and Louveaux (1988), where the recourse is disaggregated by scenarios. Another form of disaggregation is proposed for the VRPSD by Séguin (1996). Their idea is to replace θ by a set of variables and to generate an optimality cut for each route of each integer solution that is found. In their approach, each route is associated with a specific variable, and only the variables of the active routes are positive. The fact that these optimality cuts are active for all the solutions that include the same route is a significant advantage from a computational perspective, as illustrated by Côté et al. (2020). Finally, Côté et al. (2020) introduced new LBFs for the VRPSD and proposed adding them to the model each time a new path with a positive recourse is generated.

Our work extends from Séguin (1996) and Côté et al. (2020) by introducing a new approach for approximating the recourse. Similarly to Séguin (1996), θ is replaced by a variable θ_i for each customer i. This gives the following objective function.

$$\min \sum_{(i,j)\in E} c_{ij} x_{ij} + \sum_{i\in N} \theta_i \tag{50}$$

Then, a MP is constructed with the new objective function, relaxed integrality constraints, and rounded capacity inequalities. In each iteration of the method, when a feasible integer solution is found, an optimality cut is added for each route $r = (0, p, 0)$ having a positive recourse $(\mathcal{Q}(r) > 0)$, where $p = (i_1, \ldots, i_t)$. Such a path p is composed of $|p| = t - 1$ edges. For a first-stage solution x, let us denote by $x(p) = \sum_{j=1}^{t-1} x_{i_j i_{j+1}}$ the number of these edges that are active. Also, let $N(r)$ be the set of customers that are visited in path p. The new optimality cut follows.

$$\sum_{i\in N(r)} \theta_i \geq \mathcal{Q}(r)\,(x(p) - |p| + 1) \tag{51}$$

We refer to these cuts as path cuts (P-Cuts). Since they do not include the depot edges, as opposed to the route cuts of Séguin (1996), these P-Cuts are more general. In the specific case of a route $r = (0, i, 0)$ visiting a single customer i, inequality (51) reduces to $\theta_i \geq \mathcal{Q}(r)$.

We now prove that the P-Cuts are valid optimality cuts. By Laporte and Louveaux (1993), this implies that the DL-shaped method yields an optimal solution in a finite number of steps.

Proposition 11. *Let x^ν be a first-stage feasible solution and $\mathcal{Q}(x^\nu)$ the corresponding recourse cost. The set of optimality cuts (51), for $r \in \mathcal{R}^\nu$, is valid if the problem has the monotonicity property.*

Proof. Let x^λ be a solution satisfying constraints (2)-(7). For the right-hand side of (51) we have $(x^\lambda(p) - |p| + 1) \leq 1$, with equality if and only if path p is a connected subpath of a route r contained in solution x^λ. In particular, for $x^\lambda = x^\nu$, for each route $r \in \mathcal{R}^\nu$, the right-hand side of (51) takes the value $\mathcal{Q}(r)$. Together, these $|\mathcal{R}^\nu|$ constraints imply that a solution (x^ν, θ^ν) respects $\sum_{i\in N} \theta_i^\nu = \sum_{r\in\mathcal{R}^\nu} \sum_{i\in N(r)} \theta_i^\nu \geq \sum_{r\in\mathcal{R}^\nu} \mathcal{Q}(r) = \mathcal{Q}(x^\nu)$. Therefore, these cuts bind $\sum_{i\in N} \theta_i^\nu$ to $\mathcal{Q}(x^\nu)$ and the objective value of a feasible solution (x^ν, θ^ν) for the MP cannot be smaller than that of x^ν for problem (1)-(7).

Let $\theta^*(x^\lambda)$ be the minimizer of $\sum_{i\in N} \theta_i^\lambda$ over the feasible set specified by the first-stage solution x^λ and by the constraints (51) that are contained in the MP at a given iteration. Next, we prove that constraints (51) produce valid lower bounds on $\sum_{i\in N} \theta_i^\lambda$. To do so, we show that the objective value of a solution $(x^\lambda, \theta^*(x^\lambda))$ to the MP is never larger than that of x^λ for problem (1)-(7). The MP is maximally constrained if (51) has been added to the model for each possible route r. Let $\mathcal{R}^\lambda$ be the set of routes of solution x^λ. Under this maximal set of optimality cuts, for each $r \in \mathcal{R}^\lambda$, the sum $\sum_{i\in N(r)} \theta_i^\lambda$ is constrained to be larger than or equal to $F(\tilde{r})$ for each route $\tilde{r} = (0, \tilde{p}, 0)$ such that $\tilde{p}$ is a connected subpath of route r. Since it is assumed that the problem respects the monotonicity property, it follows from Proposition 3 that the most restrictive of these constraints is $\sum_{i\in N(r)} \theta_i^\lambda \geq \mathcal{Q}(r)$. Hence, by definition of $\theta^*(x^\lambda)$, $\sum_{i\in N(r)} \theta_i^*(x^\lambda) = \mathcal{Q}(r)$ and thus, $\sum_{i\in N} \theta_i^*(x^\lambda) = \sum_{r\in\mathcal{R}^\lambda} \sum_{i\in N(r)} \theta_i^*(x^\lambda) = \sum_{r\in\mathcal{R}^\lambda} \mathcal{Q}(r) = \mathcal{Q}(x^\lambda)$. The objective value of $(x^\lambda, \theta^*(x^\lambda))$ and x^λ for their respective problems are therefore identical when the MP is maximally constrained. When cuts are still to be generated, the objective value of $(x^\lambda, \theta^*(x^\lambda))$ for the MP can thus only be less than or equal to that of x^λ for problem (1)-(7). □

As shown in Laporte and Louveaux (1993), the linear relaxation of (1)-(7) can be improved by introducing a general lower bound on the recourse using a precomputed value L.

$$\sum_{i \in N} \theta_i \geq L \tag{52}$$

Unfortunately, such a bound is generally close to 0 for problems with an unlimited fleet, as using more vehicles reduces the probability of having a failure. In particular, the recourse cost is close to 0 when each vehicle only serves a single client. It can therefore be useful to compute a lower bound L_m on the recourse given that exactly $m \in M$ vehicles are being used. L_m is computed using the method defined in Section 5.3.3. This leads to the following valid lower bound inequalities.

$$\sum_{i \in N} \theta_i \geq \sum_{m \in M} L_m z_m \tag{53}$$

To conclude this section, we propose a different optimality cut that does not require the recourse to be monotonic. In the absence of monotonicity, the validity of our method can be preserved by adding the depot edges to cut (51). This way, the cut associated with a route $r = (0, i_1, \ldots, i_t, 0)$ is only active when this route appears in a first-stage solution. The coefficients of the non-depot edge variables are equal to two in order to avoid the cut being active in other solutions. The modified optimality cut follows.

$$\sum_{i \in N(r)} \theta_i \geq \mathcal{Q}(r) \left(\sum_{j=1}^{t-1} 2 x_{i_j i_{j+1}} + x_{0 i_1} + x_{0 i_t} - 2(t-1) - 1 \right) \tag{54}$$

5.2   Lower bounding functionals

This section presents a new type of LBFs that extends some ideas proposed in Côté et al. (2020). Similarly to the previously introduced optimality cuts, this type of cuts uses variables to bound the recourse for sets of customers S. The LBFs of Côté et al. (2020) use a different set of variables than those used for optimality cuts, which has the disadvantage of making the LBFs independent from the optimality cuts. This work improves over them by using the same θ_i variables for both types of cuts. The new LBFs bound the recourse using all the θ_i variables associated with a set of customers S. A cut is active whenever a path visits all the customers of this set consecutively. The new LBFs are referred to as the set cuts (S-Cuts) and are as follows:

$$\sum_{i \in S} \theta_i \geq L(S) \left(x(S) - |S| + \left\lceil \frac{\sum_{i \in S} \mu_i}{Q} \right\rceil + 1 \right), \tag{55}$$

where $x(S) = \sum_{(i,j) \in E(S)} x_{ij}$. The S-Cut (55) of a set S is based on the least possible recourse cost of visiting the customers of S. Such a lower bound $L(S)$ is valid if, for any set of feasible paths $\{p_1, \ldots, p_m\}$ forming a partition of S, the following inequality is respected.

$$L(S) \leq \sum_{j=1}^{m} \mathcal{Q}(p_j) \tag{56}$$

Proposition 12. *The S-Cut (55) is valid if the problem has the monotonicity property.*

Proof. Let x^λ be a solution satisfying constraints (2)-(7). If $x^\lambda(S) \leq |S| - \left\lceil \frac{\sum_{i \in S} \mu_i}{Q} \right\rceil - 1$, then the right-hand side of (55) is non-positive and the constraint is thus respected by any feasible non-negative solution $(x^\lambda, \theta^\lambda)$ to the MP. Otherwise, $x^\lambda(S) = |S| - \left\lceil \frac{\sum_{i \in S} \mu_i}{Q} \right\rceil$. This means that solution x^λ uses the smallest number of vehicles allowed by the capacity constraints to visit the customers of S and that, in each route $r \in \mathcal{R}^\lambda$, the customers of S are visited consecutively. In this case, the S-Cut reduces to $\sum_{i \in S} \theta_i \geq L(S)$. Let $\{p_1^\lambda, \ldots, p_m^\lambda\}$, be the set of paths forming solution x^λ, where $m = \left\lceil \frac{\sum_{i \in S} \mu_i}{Q} \right\rceil$. To prove the validity of constraint (55), we have to demonstrate that $L(S) \leq \sum_{j=1}^m \mathcal{Q}(p_j^\lambda)$. For each $j \in \{1, \ldots, m\}$, the segment of path p_j^λ that is composed of customers of S is a connected subpath p_j of p_j^λ and the set $\{p_1, \ldots, p_m\}$ of these subpaths form a partition of S. Since it is assumed that the problem respects the monotonicity property, Proposition 3 implies that $\mathcal{Q}(p_j^\lambda) \geq \mathcal{Q}(p_j) \; \forall j \in \{1, \ldots, m\}$. By definition of the lower bound $L(S)$, the sequence of inequalities $L(S) \leq \sum_{j=1}^m \mathcal{Q}(p_j) \leq \sum_{j=1}^m \mathcal{Q}(p_j^\lambda)$ allows to conclude. □

5.3 Recourse lower bounds

This section presents lower bounds on the recourse used by the S-Cuts for a given set of customers S to be served by m vehicles. Some lower bounds from the literature are first presented.

To our knowledge, the first lower bound on the recourse of a given set S is due to Laporte et al. (2002), where they suppose vehicles are only allowed to fail once. In the best case, the failures would happen for the customers of S that are the closest to the depot. Then, the sum of demands is seen as a continuous quantity that has to be distributed among the m vehicles. The problem is modeled as a non-linear program having for objective the minimization of the expected recourse.

A simpler bound, which is based on the assumption that there is one unique large vehicle with capacity mQ, is proposed by Louveaux and Salazar-González (2018). In the best case, the l-th failure would happen at the l-th customer closest to the depot. Then, the expected recourse cost can be easily calculated by summing the probability of the l-th failure multiplied by twice the l-th closest distance to the depot among the customers of S.

In the following, we propose three new bounds, $L_1(S)$, $L_2(S)$ and $L_3(S)$, which have the advantage of providing tighter values than those of Laporte et al. (2002) and Louveaux and Salazar-González (2018). All these bounds are compared later in Section 7.

5.3.1 Single-route lower bound.

The lower bound $L_1(S)$ introduced in this section applies to sets of customers that respect the monotonicity condition and can be assigned to a single vehicle. Let $\mathcal{P}^1(S)$ denote the set of all possible paths that visit the customers of S consecutively. Proposition 13 shows that sorting these customers in non-increasing order of their distance to the depot produces the path of $\mathcal{P}^1(S)$ that minimizes the recourse cost. By the monotonicity condition, the probability that a restocking trip occurs at a given customer increases with its position in the path. The idea of this bound is thus to place customers that are far from the depot at the beginning of the path to avoid costly restocking trips.

Proposition 13. *Let $S \subseteq N$ be a set of customers respecting the monotonicity condition. Let $p^* = (i_1^*, i_2^*, \ldots, i_{|S|}^*) \in \mathcal{P}^1(S)$ be a path such that $c_{0i_1^*} \geq c_{0i_2^*} \geq \cdots \geq c_{0i_{|S|}^*}$. The expected recourse $\mathcal{Q}(p^*)$ of path p^* is a valid lower bound on the expected recourse $\mathcal{Q}(p)$ of any path $p \in \mathcal{P}^1(S)$.*

Proof. Let us consider a path $p = (i_1, \ldots, i_{|S|}) \in \mathcal{P}^1(S)$ in which the customers are not sorted in non-increasing order of distances to the depot. In particular, this implies that there exists at least one index $h \in \{1, \ldots, |S|-1\}$ such that $c_{0i_h} < c_{0i_{h+1}}$. We show that the expected recourse of path $p' = (i_1, \ldots, i_{h+1}, i_h, \ldots, i_{|S|}) \in \mathcal{P}^1(S)$, in which the order of visit of customers i_h and i_{h+1} is inverted, is less than or equal to that of path p.

We first develop the expression of $\mathcal{Q}(p)$:

$$
\mathcal{Q}(p) = 2 \sum_{j=1}^{|S|} \sum_{l=1}^{\infty} \mathbb{P}\left(\sum_{k=1}^{j-1} \xi_{i_k} \leq lQ < \sum_{k=1}^{j} \xi_{i_k} \right) c_{0i_j}
$$

$$
= 2 \sum_{j=1}^{h-1} \sum_{l=1}^{\infty} \mathbb{P}\left(\sum_{k=1}^{j-1} \xi_{i_k} \leq lQ < \sum_{k=1}^{j} \xi_{i_k} \right) c_{0i_j} + 2 \sum_{l=1}^{\infty} \mathbb{P}\left(\sum_{k=1}^{h-1} \xi_{i_k} \leq lQ < \sum_{k=1}^{h} \xi_{i_k} \right) c_{0i_h}
$$

$$
+ 2 \sum_{l=1}^{\infty} \mathbb{P}\left(\sum_{k=1}^{h} \xi_{i_k} \leq lQ < \sum_{k=1}^{h+1} \xi_{i_k} \right) c_{0i_{h+1}} + 2 \sum_{j=h+2}^{|S|} \sum_{l=1}^{\infty} \mathbb{P}\left(\sum_{k=1}^{j-1} \xi_{i_k} \leq lQ < \sum_{k=1}^{j} \xi_{i_k} \right) c_{0i_j}
$$

$$
= 2 \sum_{j=1}^{h-1} \sum_{l=1}^{\infty} \mathbb{P}\left(\sum_{k=1}^{j-1} \xi_{i_k} \leq lQ < \sum_{k=1}^{j} \xi_{i_k} \right) c_{0i_j} + 2 \sum_{l=1}^{\infty} \mathbb{P}\left(\tilde{\xi} \leq lQ < \tilde{\xi} + \xi_{i_h} \right) c_{0i_h}
$$

$$
+ 2 \sum_{l=1}^{\infty} \mathbb{P}\left(\tilde{\xi} + \xi_{i_h} \leq lQ < \tilde{\xi} + \xi_{i_h} + \xi_{i_{h+1}} \right) c_{0i_{h+1}} + 2 \sum_{j=h+2}^{|S|} \sum_{l=1}^{\infty} \mathbb{P}\left(\sum_{k=1}^{j-1} \xi_{i_k} \leq lQ < \sum_{k=1}^{j} \xi_{i_k} \right) c_{0i_j},
$$

where $\tilde{\xi} = \sum_{k=1}^{h-1} \xi_{i_k}$. Doing the same for $\mathcal{Q}(p')$, we obtain:

$$
\mathcal{Q}(p') = 2 \sum_{j=1}^{h-1} \sum_{l=1}^{\infty} \mathbb{P}\left(\sum_{k=1}^{j-1} \xi_{i_k} \leq lQ < \sum_{k=1}^{j} \xi_{i_k} \right) c_{0i_j} + 2 \sum_{l=1}^{\infty} \mathbb{P}\left(\tilde{\xi} \leq lQ < \tilde{\xi} + \xi_{i_{h+1}} \right) c_{0i_{h+1}}
$$

$$
+ 2 \sum_{l=1}^{\infty} \mathbb{P}\left(\tilde{\xi} + \xi_{i_{h+1}} \leq lQ < \tilde{\xi} + \xi_{i_{h+1}} + \xi_{i_h} \right) c_{0i_h} + 2 \sum_{j=h+2}^{|S|} \sum_{l=1}^{\infty} \mathbb{P}\left(\sum_{k=1}^{j-1} \xi_{i_k} \leq lQ < \sum_{k=1}^{j} \xi_{i_k} \right) c_{0i_j}
$$

Since the expected recourse cost at customers i_1 to i_{h-1} and at customers i_{h+2} to $i_{|S|}$ is identical in both paths, the inequality $\mathcal{Q}(p) \geq \mathcal{Q}(p')$ simplifies to:

$$
\mathcal{Q}(p) \geq \mathcal{Q}(p')
$$

$$
\Longleftrightarrow 2 \sum_{l=1}^{\infty} \mathbb{P}\left(\tilde{\xi} \leq lQ < \tilde{\xi} + \xi_{i_h} \right) c_{0i_h} + 2 \sum_{l=1}^{\infty} \mathbb{P}\left(\tilde{\xi} + \xi_{i_h} \leq lQ < \tilde{\xi} + \xi_{i_h} + \xi_{i_{h+1}} \right) c_{0i_{h+1}} \geq
$$

$$
2 \sum_{l=1}^{\infty} \mathbb{P}\left(\tilde{\xi} \leq lQ < \tilde{\xi} + \xi_{i_{h+1}} \right) c_{0i_{h+1}} + 2 \sum_{l=1}^{\infty} \mathbb{P}\left(\tilde{\xi} + \xi_{i_{h+1}} \leq lQ < \tilde{\xi} + \xi_{i_{h+1}} + \xi_{i_h} \right) c_{0i_h}
$$

$$
\Longleftrightarrow \sum_{l=1}^{\infty} \left(\mathbb{P}\left(\tilde{\xi} \leq lQ < \tilde{\xi} + \xi_{i_h} \right) - \mathbb{P}\left(\tilde{\xi} + \xi_{i_{h+1}} \leq lQ < \tilde{\xi} + \xi_{i_{h+1}} + \xi_{i_h} \right) \right) c_{0i_h} \geq
$$

$$
\sum_{l=1}^{\infty} \left(\mathbb{P}\left(\tilde{\xi} \leq lQ < \tilde{\xi} + \xi_{i_{h+1}} \right) - \mathbb{P}\left(\tilde{\xi} + \xi_{i_h} \leq lQ < \tilde{\xi} + \xi_{i_h} + \xi_{i_{h+1}} \right) \right) c_{0i_{h+1}}
$$

$$\Longleftrightarrow \sum_{l=1}^{\infty} \left(\mathbb{P} \left(\tilde{\xi} + \xi_{i_h} > lQ \right) - \mathbb{P} \left(\tilde{\xi} > lQ \right) - \mathbb{P} \left(\tilde{\xi} + \xi_{i_{h+1}} + \xi_{i_h} > lQ \right) + \mathbb{P} \left(\tilde{\xi} + \xi_{i_{h+1}} > lQ \right) \right) c_{0i_h} \geq$$

$$\sum_{l=1}^{\infty} \left(\mathbb{P} \left(\tilde{\xi} + \xi_{i_{h+1}} > lQ \right) - \mathbb{P} \left(\tilde{\xi} > lQ \right) - \mathbb{P} \left(\tilde{\xi} + \xi_{i_h} + \xi_{i_{h+1}} > lQ \right) + \mathbb{P} \left(\tilde{\xi} + \xi_{i_h} > lQ \right) \right) c_{0i_{h+1}}$$

$$\Longleftrightarrow \sum_{l=1}^{\infty} \left(-\mathbb{P} \left(\tilde{\xi} + \xi_{i_h} \leq lQ \right) + \mathbb{P} \left(\tilde{\xi} \leq lQ \right) + \mathbb{P} \left(\tilde{\xi} + \xi_{i_{h+1}} + \xi_{i_h} \leq lQ \right) - \mathbb{P} \left(\tilde{\xi} + \xi_{i_{h+1}} \leq lQ \right) \right) c_{0i_h} \geq$$

$$\sum_{l=1}^{\infty} \left(-\mathbb{P} \left(\tilde{\xi} + \xi_{i_{h+1}} \leq lQ \right) + \mathbb{P} \left(\tilde{\xi} \leq lQ \right) + \mathbb{P} \left(\tilde{\xi} + \xi_{i_h} + \xi_{i_{h+1}} \leq lQ \right) - \mathbb{P} \left(\tilde{\xi} + \xi_{i_h} \leq lQ \right) \right) c_{0i_{h+1}}$$

$$\Longleftrightarrow (c_{0i_{h+1}} - c_{0i_h}) \sum_{l=1}^{\infty} \left(\mathbb{P} \left(\tilde{\xi} + \xi_{i_h} \leq lQ \right) - \mathbb{P} \left(\tilde{\xi} \leq lQ \right) - \mathbb{P} \left(\tilde{\xi} + \xi_{i_{h+1}} + \xi_{i_h} \leq lQ \right) + \mathbb{P} \left(\tilde{\xi} + \xi_{i_{h+1}} \leq lQ \right) \right) \geq 0$$

$$\Longleftarrow \mathbb{P} \left(\tilde{\xi} + \xi_{i_h} \leq lQ \right) - \mathbb{P} \left(\tilde{\xi} + \xi_{i_{h+1}} + \xi_{i_h} \leq lQ \right) \geq \mathbb{P} \left(\tilde{\xi} \leq lQ \right) - \mathbb{P} \left(\tilde{\xi} + \xi_{i_{h+1}} \leq lQ \right) \qquad , \forall l \geq 1$$

Since S respects the monotonicity condition, the last inequality, which corresponds to inequality (10) with $\tilde{S} = \{i_1, \ldots, i_{h-1}\}$ and $(a, b) = (i_h, i_{h+1})$, is verified for any $l \in \mathbb{N}$.

This implies that, starting from any path $p \in \mathcal{P}^1(S)$, one can iteratively perform inversions of consecutive customers $(a, b) \in S \times S$ such that $c_a < c_b$ until all the customers are sorted in non-increasing order of their distance to the depot without increasing the expected recourse cost. Therefore, the set $\arg\min_{p \in \mathcal{P}^1(S)} \mathcal{Q}(p)$ contains at least one path $\tilde{p} = (\tilde{i}_1, \ldots, \tilde{i}_{|S|})$ such that $c_{0\tilde{i}_1} \geq c_{0\tilde{i}_2} \geq \cdots \geq c_{0\tilde{i}_{|S|}}$.

To conclude that $p^* \in \arg\min_{p \in \mathcal{P}^1(S)} \mathcal{Q}(p)$, and thus that $\mathcal{Q}(p^*)$ is a valid lower bound on the expected recourse of any path $p \in \mathcal{P}^1(S)$, it remains to show that $\mathcal{Q}(\tilde{p}) = \mathcal{Q}(p^*)$. To do so, we sort the distances separating the customers of S from the depot and remove the duplicates to obtain an ordered list $c_{0(1)} > \cdots > c_{0(t)}$ of length $t \leq |S|$. For each $s \in \{1, \ldots, t\}$, let us denote by $C_s = \{i \in S : c_{0i} = c_{0(s)}\}$ the set of customers that share the same distance $c_{0(s)}$ from the depot and by $\xi^s = \sum_{i \in C_s} \xi_i$ the sum of their demands. By construction, in both paths $\tilde{p}$ and p^*, a permutation of C_1 will first be visited, followed by a permutation of C_2 and so on until C_t. Denoting the number of customers in group C_s as $v_s = |C_s|$ and the total number of customers visited before moving to group C_{s+1} as $V_s = \sum_{q=1}^{s} v_r$, we can demonstrate the expected equality.

$$\mathcal{Q}(\tilde{p}) = 2 \sum_{j=1}^{|S|} \sum_{l=1}^{\infty} \mathbb{P} \left(\sum_{k=1}^{j-1} \xi_{\tilde{i}_k} \leq lQ < \sum_{k=1}^{j} \xi_{\tilde{i}_k} \right) c_{0\tilde{i}_j} \qquad (57)$$

$$= 2 \sum_{s=1}^{t} \sum_{j=V_{s-1}+1}^{V_s} \sum_{l=1}^{\infty} \mathbb{P} \left(\sum_{k=1}^{j-1} \xi_{\tilde{i}_k} \leq lQ < \sum_{k=1}^{j} \xi_{\tilde{i}_k} \right) c_{0(s)} \qquad (58)$$

$$= 2 \sum_{s=1}^{t} c_{0(s)} \sum_{l=1}^{\infty} \left(\sum_{j=V_{s-1}+1}^{V_s} \mathbb{P} \left(\sum_{k=1}^{j-1} \xi_{\tilde{i}_k} \leq lQ < \sum_{k=1}^{j} \xi_{\tilde{i}_k} \right) \right) \qquad (59)$$

$$= 2 \sum_{s=1}^{t} c_{0(s)} \sum_{l=1}^{\infty} \mathbb{P} \left(\sum_{q=1}^{s-1} \xi^q \leq lQ < \sum_{q=1}^{s} \xi^q \right) \qquad (60)$$

$$= 2 \sum_{s=1}^{t} c_{0(s)} \sum_{l=1}^{\infty} \left(\sum_{j=V_{s-1}+1}^{V_s} \mathbb{P} \left(\sum_{k=1}^{j-1} \xi_{i_k^*} \le lQ < \sum_{k=1}^{j} \xi_{i_k^*} \right) \right) \tag{61}$$

$$= 2 \sum_{s=1}^{t} \sum_{j=V_{s-1}+1}^{V_s} \sum_{l=1}^{\infty} \mathbb{P} \left(\sum_{k=1}^{j-1} \xi_{i_k^*} \le lQ < \sum_{k=1}^{j} \xi_{i_k^*} \right) c_{0(s)} \tag{62}$$

$$= 2 \sum_{j=1}^{|S|} \sum_{l=1}^{\infty} \mathbb{P} \left(\sum_{k=1}^{j-1} \xi_{i_k^*} \le lQ < \sum_{k=1}^{j} \xi_{i_k^*} \right) c_{0i_j^*} \tag{63}$$

$$= \mathcal{Q}(p^*) \tag{64}$$

The equivalence of equalities (59) and (60) and, analogously, of equalities (60) and (61), comes from the fact that the l-th restocking trip will occur at a customer of set C_s if and only if the total demand of the customers of sets C_1 to C_{s-1} is less than or equal to lQ and the total demand of the customers of sets C_1 to C_s exceeds lQ. The sum of the probabilities of the mutually exclusive events that are considered for each $s \in \{1, \dots, t\}$ and for each $l \in \mathbb{N}$ in both equations (59) and (61) is thus equal to the probability of their disjunction, which is given in (60). $\square$

The definition of the single-route lower bound follows.

$$L_1(S) = \mathcal{Q}(p^*) \tag{65}$$

5.3.2 Dynamic programming algorithm.

The lower bound proposed in this section applies to sets of customers S that can be assigned to $m \ge 2$ vehicles. It is inspired by the lower bound of Laporte et al. (2002), where the demands are seen as continuous quantities that must be distributed among the vehicles. This new bound, which requires the monotonicity of the recourse function, improves over Laporte et al. (2002) in two ways. First, the value of the lower bound can be greatly increased by considering that multiple failures can happen for a vehicle. This is done by computing for each vehicle k and for each quantity of expected demand $q \in [0, Q]$ the least-cost sequence of customers by taking advantage of Proposition 13. The second improvement is that our search for an assignment of the demands can be restricted to integer solutions through the use of a new dynamic programming algorithm. The new bound is computed in two steps.

In the first step, the least-cost sequence for assigning $q_h \in [0, Q]$ units of expected demand to a vehicle using the $|S| - h + 1$ farthest customers of S is computed. To do so, the customers of S are sorted by non-decreasing distance to the depot. The least-cost sequence is computed by dynamic programming over a set of $|S|$ stages indexed by $h = 1$ to $|S|$. The state q_h represents the number of units of expected demand assigned in stage h. In state q_h, we denote the total demand of the assigned customers by the RV ξ_{q_h}. Its distribution is given by $\xi_{q_h} \sim \text{Poisson}(q_h)$ if the demands are modeled as Poisson variables and by $\xi_{q_h} \sim \mathcal{N}(q_h, var(q_h))$ if the demands are normally distributed. In the case of normal demands, we set $var(q_h) = \min\{\sum_{i=h}^{|S|} \sigma_i^2 z_i \mid \sum_{i=h}^{|S|} \mu_i z_i = q_h, z_i \in \{0, 1\}\}$, which is the smallest possible variance for the total demand of the assigned customers. Function

$G_h(q_h)$ finds the value of the least-cost sequence.

$$G_h(q_h) = \begin{cases} G_{h+1}(q_h), & \text{for } h < |S| \text{ and } q_h < \mu_h, \\ \min\left\{G_{h+1}(q_h), G_{h+1}(q_h - \mu_h) + \sum_{l=1}^{\infty} P(\xi_{q_h - \mu_h} \leq lQ < \xi_{q_h})\right\}, & \text{for } h < |S| \text{ and } q_h \geq \mu_h, \\ 1 - P(\xi_{q_h} \leq Q), & \text{for } h = |S| \text{ and } q_h = \mu_h, \\ \infty, & \text{otherwise.} \end{cases}$$

The first two equations select the lowest-cost assignment between the previous stage in state q_h and the previous stage in state $q_h - \mu_h$ plus the probability of failure when customer h is added when there are already $q_h - \mu_h$ units on board. The third equation is for the last customer. The last one manages the initial cases.

In the second step, the bound is obtained by computing the least-cost assignment of the $\sum_{i \in S} \mu_i$ units of expected demand to the vehicles. This is done by defining another dynamic program that uses the previously computed functions $G_h(q)$ and only allows the h-th customer to be assigned to vehicle k if $k \leq h$. This dynamic program is defined over m stages indexed by $k = 1$ to m. The state q_k represents the number of units of expected demand that are assigned to the first k vehicles. At stage k, variable x_k decides which quantity is assigned to vehicle k. Then, function $H_k(q_k)$ finds the optimal assignment of q_k units to the first k vehicles. It is defined as follows.

$$H_k(q_k) = \begin{cases} \infty, & \text{if } q_k > kQ, \\ G_k(q_k), & \text{for } k = 1 \text{ and } q_k \leq Q, \\ \min_{0 \leq x_k \leq \min(q_k, Q)} \{G_k(x_k) + H_{k-1}(q_k - x_k)\}, & \text{if } 2 \leq k \leq m. \end{cases}$$

Finally, the expected recourse cost of the optimal assignment of the demands is given by:

$$L_2(S) = H_m\left(\sum_{i \in S} \mu_i\right), \tag{66}$$

which can be computed in $O(m \sum_{i \in N} \mu_i)$.

5.3.3 A set-covering formulation.

The last lower bound is based on a set-covering formulation, where the linear relaxation is solved by column generation. Each column represents a path that visits some customers of S and respects the expected capacity constraint. The problem is to find a set of feasible paths of cardinality m that visits all customers of S and minimizes the total expected recourse cost. Let $\mathcal{P}(S)$ be the set of all feasible paths visiting only customers of S. We define z_p as a binary variable indicating whether path $p \in \mathcal{P}(S)$ is taken or not. Also, the binary coefficient b_{jp} indicates whether customer j is in path p. The formulation is as follows:

$$L_3(S) = \min \sum_{p \in \mathcal{P}(S)} \mathcal{Q}(p) z_p \tag{67}$$

$$\text{s.t.} \sum_{p \in \mathcal{P}(S)} b_{ip} z_p \geq 1, \qquad i \in S, \tag{68}$$

$$\sum_{p \in \mathcal{P}(S)} z_p = m, \tag{69}$$

$$z_p \in \{0, 1\}, \qquad\qquad p \in \mathcal{P}(S). \tag{70}$$

The objective function (67) minimizes the sum of the expected recourse costs of the selected paths. Constraints (68) state that each customer of S must be visited. Constraint (69) ensures that exactly m paths are selected, and constraints (70) define the domain of the variables.

As the set of paths $\mathcal{P}(S)$ is too large to be generated up front, we rely on column generation to solve the problem. An initial set of columns is first generated to obtain a feasible solution for the linear relaxation. Then, a pricing problem is solved to find a column having a negative reduced cost. If such a column is found, it is added to the model, and the linear relaxation is solved again. This process continues until no more negative reduced cost columns can be found.

The pricing problem of finding the least-cost path can be formulated using the dynamic program of the previous section. The second equation, when adding customer h, is modified by subtracting the value of the dual variable associated with the constraint (68) of customer h. A negative cost column is found if the value of the least-cost sequence is less than the value of the dual variable associated with constraint (69).

6   Implementation of the DL-shaped method

This section details the steps performed when executing the DL-shaped method.

1. Compute $L_m = L_3(N)$ for each number of vehicles $m \in M$.

2. Build model (2)-(7) with the objective function (50) and add constraint (53).

3. Solve the linear relaxation of the model.

4. Verify if there are any violated subtour or capacity constraints (4) using the CVRPSEP package of Lysgaard (2003). Add the violated inequalities (4) to the model. For each identified set S, check if its associated S-Cut is violated, and if so, add the inequality to the model.

5. If the solution is fractional, enumerate its connected components. For each component, verify if an S-Cut is violated, and if one is found, add the S-Cut to the model. If the number of nodes in the component equals the number of non-zero valued edges plus one, then the component is a path. Then, check if the P-Cut associated with the path is violated, and if so, add it to the model.

6. Go back to Step 3 if violated inequalities were found in Steps 4 or 5. Otherwise, branch on a fractional edge variable if the solution is fractional and return to Step 3.

7. For each route in the solution, add the violated P&S-Cuts to the model. Return to Step 3 if violated inequalities were found.

The procedure stops when all the nodes of the brand-and-bound tree have been explored. The incumbent solution is then optimal. In our implementation, the S-Cut associated with a set of customers S is generated using the lower bound $L_1(S)$ if the sum of their expected demands does not exceed the vehicle capacity. Otherwise, the lower bound $L_2(S)$ is used, with $m = \lceil (\sum_{i \in S} \mu_i)/Q \rceil$.

To compute the P&S-Cuts of Step 7, we iterate on the routes of the current solution and generate all the subpaths in which at most five customers are removed from the original route. The recourse cost and the single-route lower bound L_1 are then computed for each subpath. The five most violated P-Cuts and the five most violated S-Cuts are added to the model.

7 Computational experiments

This section analyzes the performance of the DL-shaped method through computational experiments. All the results were produced with a C++ implementation with Cplex 12.10 on a machine with an AMD Rome 7532 @ 2.40GHz CPU. For all the computational experiments, we defined a maximum time of one hour per instance. A quick upper bound is also provided to the solver by applying the adaptive large neighborhood search of Ropke and Pisinger (2006). The detailed results can be found at .

Section 7.1 describes the existing instances that are used in our experiments and introduces a new set of instances. In Section 7.2, the new lower bounds on the recourse are compared with classical bounds from the literature. Sections 7.3 compare our algorithm to previous methods on two different instance sets. Finally, Section 7.4 presents the performance of our method on the new set of instances and summarizes the key takeaways of our experiments.

7.1 Characterization of the instances

Computational experiments are based on three sets of instances: two from the literature and a new one. The first set is taken from Jabali et al. (2014) and contains 270 instances. In each instance, n ranges from 40 to 80, and the demands are normally distributed, with parameters (μ_i, σ_i) such that $\mu_i = 3\sigma_i$ and $\mu_i \in \{1, \ldots, 10\}$. The number of vehicles is fixed to $\bar{m} \in \{2, 3, 4\}$ and the filling coefficient $\bar{f} = \frac{\sum_{i \in N} \mu_i}{\bar{m}Q}$ ranges from 0.80 to 0.95. The DL-shaped method can be used on this set since they respect the monotonicity property. We verified that inequality (11) holds for each capacity Q in this set of instances, for each number of restocking trips $l \in \{1, 2, 3\}$, and for each possible distribution of $\xi_a \sim \mathcal{N}(\mu_a, \sigma_a^2)$, $\xi_b \sim \mathcal{N}(\mu_b, \sigma_b^2)$, and $\tilde{\xi} \sim \mathcal{N}(\tilde{\mu}, \tilde{\sigma}^2)$. We ignored the case $l \geq 4$ since the probability of more than three restocking trips to occur is negligible.

The second set is taken from classical CVRP instances and contains 83 instances from the A, B, E, M, and P sets described in Gauvin et al. (2014). In these instances, the customer demands are modeled as Poisson RVs. Regarding customer demands, the parameter μ_i is set to the deterministic demand for customer i of the CVRP instance. All these instances have unlimited fleets.

For the new set, we generated 1980 instances using the method of Jabali et al. (2014). To obtain a challenging set of instances, we produced ten instances for each combination of the following problem parameters: $n \in \{20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120\}$, $\bar{m} \in \{2, 3, 4, 5, 6, 7\}$, and $\bar{f} \in \{0.85, 0.90, 0.95\}$. The coefficient of dispersion of the demand of each customer $i \in N$ is given by $D = \sigma_i^2/\mu_i = 1$. By Proposition 5, these instances therefore respect the monotonicity property.

7.2 Lower bounds on the recourse

This section compares our new lower bounds on the recourse to those previously used in the literature based on the relative gap they provide when applied to an optimal first-stage solution. This

measure is defined as follows.

$$\text{Gap\%} = \frac{\text{Recourse in the best or optimal solution} - \text{lower bound}}{\text{Recourse in the best or optimal Solution}} \times 100$$

For the first set of instances of Jabali et al. (2014), the results are grouped by number of customers, whereas the results are presented separately for each set A, B, E, M, and P for the CVRP instances. In Tables 1 and 2, the columns LLVH02 and LSG18 present the gaps obtained from the implementation of the lower bound from Laporte et al. (2002) and Louveaux and Salazar-González (2018), respectively. The performance of our three lower bounds, L_1, L_2, and L_3 is also reported. Column L_1 presents the gaps obtained from computing $\sum_{r \in \mathcal{R}^*} L_1(r)$, where $\mathcal{R}^*$ is the set of routes obtained in either the optimal or best solution found at the end of the DL-shaped method. Columns L_2 and L_3 are obtained by computing $L_2(N)$ and $L_3(N)$. The average computing time of $L_3(N)$ is also reported. The computing time of all the other bounds is negligible.

n	#	LLVH02	LSG18	L_1	L_2	L_3	L_3 (s)
40	30	99.9	99.9	3.5	89.7	60.0	0.9
50	60	96.0	99.6	6.9	85.1	55.1	2.9
60	90	88.6	97.2	7.4	78.5	50.8	6.4
70	60	86.3	96.7	9.0	73.8	44.2	13.7
80	30	76.6	94.5	10.2	61.4	30.4	21.7
Avg.		89.5	97.6	7.4	77.7	48.1	8.8

Table 1: Relative gap of lower bounds on the instances of Jabali et al. (2014)

Set	#	LLVH02	LSG18	L_1	L_2	L_3	L_3 (s)
A	26	68.8	100.0	1.7	68.1	66.1	0.6
B	23	56.5	99.9	0.5	55.9	53.7	0.7
E	11	64.0	99.5	1.8	63.3	61.8	2.0
M	1	55.2	99.0	6.2	52.4	42.6	2.7
P	22	37.5	95.2	1.6	36.3	34.4	0.6
Avg.		56.3	98.6	1.4	55.5	53.4	0.8

Table 2: Relative gap of lower bounds on the CVRP instances

The lower bound of Louveaux and Salazar-González (2018) is systematically the weakest, followed by that of Laporte et al. (2002). Their average gap increase with both the fleet size and the number of customers. The bound of Louveaux and Salazar-González (2018) assumes a single large vehicle and therefore becomes a poor approximation of the recourse as the fleet size increases. Regarding the bound of Laporte et al. (2002), it suffers from an underestimation of the variance of the total demand on a route, which worsens as the number of customers grows.

These results demonstrate that our new bounds significantly improve over the previous bounds from the literature. The set covering bound L_3 achieves an average gap of approximately 50% on average on both sets of instances. The DP bound L_2 is generally slightly less tight than L_3 but is significantly cheaper to compute. The good computational performance of L_2 can be attributed to the fact that the least-cost sequence can be computed by sorting the customers according to their distance to the depot. The single-route lower bound L_1 is by far the tightest, with an average

gap of 7.4% for the first set and 1.4% for the second. For a given set of customers, the quality of bounds L_1 and L_2 directly depends on the similarity of the least-cost sequence for the recourse cost provided by Proposition 13 and the total least-cost sequence that takes the first-stage routing cost into account. In particular, the performance of bound L_1 reported in Table 2 indicates the substantial similarity of these sequences for the CVRP instances.

7.3 Results for the instances of the literature

This section reports results on the two sets of instances of the literature. It compares the efficiency of the DL-shaped method against other B&C methods and against B&P methods.

Table 3 presents the results against the integer L-shaped methods of Jabali et al. (2014) and Hoogendoorn and Spliet (2022). The performance of the three methods is presented under the columns JRGL14, HS22, and DL-Shaped of Table 3. Each row of Table 3 contains 30 instances with the same number of customers n and the same fleet size $\bar{m}$. The number of optimal solutions found by each method and their average computing time are reported, as well as the average optimality gap of the unsolved instances.

n	$\bar{m}$	JRGL14			HS22			DL-Shaped	
		Opt	Sec	Gap%	Opt	Sec	Gap%	Opt	Sec
40	4	9	1240.7	1.5	28	91.9	1.6	30	2.0
50	3	16	6918.0	0.7	29	101.5	2.9	30	9.8
50	4	5	1360.8	1.9	25	224.8	1.7	30	30.9
60	2	24	1393.0	0.4	30	72.2	-	30	1.2
60	3	6	2766.0	0.7	27	191.4	1.2	30	16.7
60	4	3	4922.0	2.0	25	262.1	1.9	30	25.3
70	2	17	2577.5	0.5	30	99.2	-	30	2.0
70	3	9	1753.3	1.5	24	563.7	1.5	30	46.2
80	2	13	1809.2	0.5	28	193.7	1.0	30	19.8
Total or avg.		102	2711.4	1.2	246	185.7	1.6	270	17.1

Table 3: Performance comparison with Jabali et al. (2014); Hoogendoorn and Spliet (2022) on the instances of Jabali et al. (2014)

Table 3 indicates that the DL-shaped method achieves state-of-the-art results by solving to optimality all the 270 instances of the first set, while Jabali et al. (2014) and Hoogendoorn and Spliet (2022) respectively solve 102 and 246 instances. Furthermore, our average resolution time is 17.1 seconds per instance, compared to 2711.4 and 185.72 seconds for the instances that were solved to optimality by Jabali et al. (2014) and Hoogendoorn and Spliet (2022), respectively. These results highlight the advantage of disaggregating the recourse in the context of the integer L-shaped method, as well as the capacity of our LBFs to approximate the recourse function efficiently.

In Table 4, the DL-shaped method is compared to the algorithms of Christiansen and Lysgaard (2007), Gauvin et al. (2014), and Florio et al. (2020) on the second set of instances. These three methods, which all implement a B&P approach, are denoted by CL07, GDG14, and FHM20. Note that some entries are missing for CL07 and FHM20, as they did not report results for all instances.

Set	#	CL07	GDG14	FHM20	DL-Shaped	Gap%	Time (s)
A	26	6/19	22	14/27	9	12.4	292
B	23	-	7	-	6	11.8	930
E	11	2/3	5	3/8	5	9.0	213
M	1	-	1	-	-	5.9	-
P	22	11/18	20	17	10	7.8	183
Total or avg.	83	19/40	55	34/52	30	10.7	370

Table 4: Performance comparison with Christiansen and Lysgaard (2007); Gauvin et al. (2014); Florio et al. (2020) on the CVRP instances

The DL-shaped method can solve more than a third of the 83 instances of this set. This is a new milestone for a B&C approach, as the previous algorithms from the literature have not been successful in solving these CVRP instances due to their tight capacity constraints. Our method achieves comparable performance as GDG14 for the B and E sets and generally produces solutions of good quality, with the average gaps being close to 10% for all sets.

7.4 Results for the new instances

This section presents the results on our new set of instances. For each number of customers and vehicles, Table 5 reports the number of instances, out of 30, that could be solved within the time limit and their average resolution time. In total, 1175 instances were solved, and the average optimality gap of the 805 unsolved instances was less than 5% in each group. Even for as many as 120 customers, instances with two vehicles were generally solved easily, with an average time of one minute per solved instance. Although the difficulty of the new instances appears to increase with the number of vehicles, some of the instances that were solved by the DL-shaped method are significantly larger than those previously found in the literature. These include two instances with 50 customers and seven vehicles and one instance with 120 customers and five vehicles.

	Number of optimal solutions						Average computing times (s)					
$n \backslash \bar{m}$	2	3	4	5	6	7	2	3	4	5	6	7
20	30	30	30	30	30	30	0.0	0.4	1.3	7.9	32.4	10.7
30	30	30	30	30	24	20	0.1	0.6	5.6	276.1	807.3	863.9
40	30	30	30	28	13	3	0.3	1.4	42.2	498.0	867.0	1718.3
50	30	30	28	23	4	2	0.6	45.1	232.2	516.1	913.6	1297.1
60	30	30	24	17	3	0	3.9	12.3	275.3	1114.8	1926.9	-
70	30	30	22	7	1	0	1.6	286.1	272.3	1653.7	400.0	-
80	30	30	19	10	2	0	17.7	198.1	1024.2	1455.7	1667.5	-
90	30	28	19	6	1	0	15.8	233.3	710.4	611.6	2184.9	-
100	30	25	18	6	0	0	16.3	396.9	956.0	1382.5	-	-
110	30	26	9	5	0	0	101.9	520.4	234.7	1484.3	-	-
120	29	21	11	1	0	0	62.2	306.3	317.4	1468.2	-	-

Table 5: Solved instances and computation times of the DL-shaped method for the new set

Figure 1 summarizes the parameters of the largest instances that could be solved and the smallest instances that remained unsolved by different B&C methods from the literature. For each number of

customers and each number of vehicles, the highest filling coefficient $\bar{f}$ for which instances have been solved and the smallest filling coefficient leading to instances that could not be solved are reported. To increase readability, we only present what we consider to be the most important configurations per paper. We also indicate whether the customer demands were modeled as deterministic (D), Poisson RVs (P), or normal RVs (N) in each set of instances. Results from Jabali et al. (2014) are omitted because they are slightly inferior to those of Hoogendoorn and Spliet (2022).
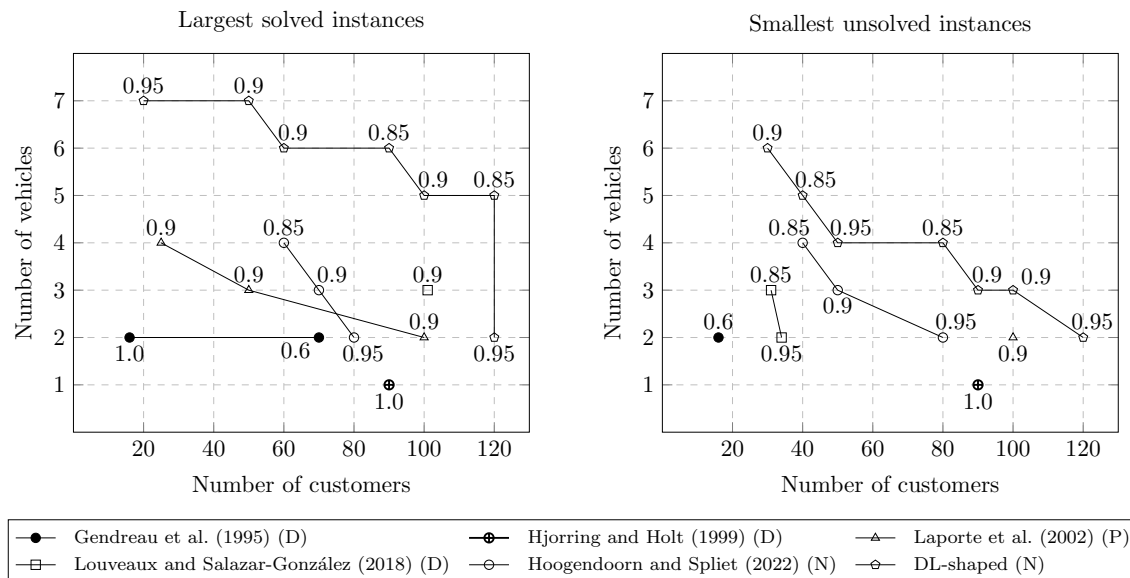

Figure 1: Largest solved and smallest unsolved instances in the literature with their filling coefficient

Although the figure should be taken with a grain of salt, these results indicate that the DL-shaped method constitutes a significant advancement in our ability to solve instances with a high number of customers and vehicles and a high filling coefficient. Although previous methods from the literature can solve some instances with a relatively high number of customers, they also fail to solve instances with as few as 30 or 40 customers.

8   Conclusion

This paper presents a new approach for solving a class of stochastic integer programs in which the first-stage solutions can be decomposed into disjoint components. The method is applied to the vehicle routing problem with stochastic demands under the detour-to-depot recourse policy. Our computational experiments show that it achieves state-of-the-art results on instances from the literature. Our approach is based on new optimality cuts and exploits new lower bounds on the recourse that are used by new lower bounding functionals. Computational experiments show that our lower bounds improve over the existing ones from the literature.

Regarding future lines of research, the method could be applied to other variants of the stochastic vehicle routing problem and other two-stage stochastic integer programs. Since the efficiency of

this method depends on the availability of tight bounds on disjoint components of the recourse, this work may motivate the derivation of such bounds for other problems of interest. Finally, as our method requires the recourse function to be monotonic, it would be interesting to conduct a systematic study of this property for other two-stage stochastic programs.

Acknowledgments

References

Ak, A. and Erera, A. L. (2007). A paired-vehicle recourse strategy for the vehicle-routing problem with stochastic demands. *Transportation Science*, 41(2):222–237.

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

Bertsimas, D., Jaillet, P., and Odoni, A. (1990). A priori optimization. *Operations Research*, 38(6):1019–1033.

Birge, J. and Louveaux, F. (1988). A multicut algorithm for two-stage stochastic linear programs. *European Journal of Operational Research*, 34(3):384–392.

Birge, J. R. and Wets, R. J.-B. (1986). *Designing approximation schemes for stochastic optimization problems, in particular for stochastic programs with recourse*, pages 54–102. Springer Berlin Heidelberg, Berlin, Heidelberg.

Christiansen, C. and Lysgaard, J. (2007). A branch-and-price algorithm for the capacitated vehicle routing problem with stochastic demands. *Operations Research Letters*, 35(6):773–781.

Côté, J.-F., Gendreau, M., and Potvin, J.-Y. (2020). The vehicle routing problem with stochastic two-dimensional items. *Transportation Science*, 54(2):453–469.

Denton, B., Miller, A., Balasubramanian, H., and Huschka, T. (2010). Optimal allocation of surgery blocks to operating rooms under uncertainty. *Operations Research*, 58(4-part-1):802–816.

Dror, M. and Ball, M. (1987). Inventory/routing: Reduction from an annual to a short-period problem. *Naval Research Logistics*, 34(6):891–905.

Dror, M., Laporte, G., and Trudeau, P. (1989). Vehicle routing with stochastic demands: Properties and solution frameworks. *Transportation science*, 23(3):166–176.

Elçi, O. and Hooker, J. (2022). Stochastic planning and scheduling with logic-based benders decomposition. *INFORMS Journal on Computing*, 34(5):2428–2442.

Florio, A., Hartl, R., and Minner, S. (2020). New exact algorithm for the vehicle routing problem with stochastic demands. *Transportation Science*, 54(4):1073–1090.

Florio, A. M., Feillet, D., Poggi, M., and Vidal, T. (2022). Vehicle routing with stochastic demands and partial reoptimization. *Transportation Science*, 56(5):1393–1408.

Gauvin, C., Desaulniers, G., and Gendreau, M. (2014). A branch-cut-and-price algorithm for the vehicle routing problem with stochastic demands. *Computers & Operations Research*, 50:141–153.

Gendreau, M., Jabali, O., and Rei, W. (2014). Chapter 8: Stochastic vehicle routing problems. In *Vehicle Routing: Problems, Methods, and Applications, Second Edition*, pages 213–239. SIAM.

Gendreau, M., Jabali, O., and Rei, W. (2016). 50th anniversary invited article—future research directions in stochastic vehicle routing. *Transportation Science*, 50(4):1163–1173.

Gendreau, M., Laporte, G., and Séguin, R. (1995). An exact algorithm for the vehicle routing problem with stochastic demands and customers. *Transportation science*, 29(2):143–155.

Hjorring, C. and Holt, J. (1999). New optimality cuts for a single-vehicle stochastic routing problem. *Annals of Operations Research*, 86:569–584.

Hoogendoorn, Y. and Spliet, R. (2022). *An improved integer L-shaped method for the vehicle routing problem with stochastic demands*. EI report reeks.

Hooker, J. and Ottosson, G. (2003). Logic-based benders decomposition. *Mathematical Programming*, 96(1):33–60.

Jabali, O., Rei, W., Gendreau, M., and Laporte, G. (2014). Partial-route inequalities for the multi-vehicle routing problem with stochastic demands. *Discrete Applied Mathematics*, 177:121–136.

Kreimer, J. and Dror, M. (1990). The monotonicity of the threshold detection probability in a stochastic accumulation process. *Computers & Operations Research*, 17(1):63–71.

Laporte, G. and Louveaux, F. (1993). The integer L-shaped method for stochastic integer programs with complete recourse. *Operations research letters*, 13(3):133–142.

Laporte, G., Louveaux, F., and Van Hamme, L. (2002). An integer L-shaped algorithm for the capacitated vehicle routing problem with stochastic demands. *Operations Research*, 50(3):415–423.

Li, Y., Côté, J.-F., Callegari-Coelho, L., and Wu, P. (2022). Novel formulations and logic-based benders decomposition for the integrated parallel machine scheduling and location problem. *INFORMS Journal on Computing*, 34(2):1048–1069.

Louveaux, F. and Salazar-González, J.-J. (2018). Exact approach for the vehicle routing problem with stochastic demands and preventive returns. *Transportation Science*, 52(6):1463–1478.

Lysgaard, J. (2003). Cvrpsep: A package of separation routines for the capacitated vehicle routing problem. Technical report, Department of Management Science and Logistics, Aarhus School of Business.

Pillac, V., Gendreau, M., Guéret, C., and Medaglia, A. (2013). A review of dynamic vehicle routing problems. *European Journal of Operational Research*, 225(1):1–11.

Rahmaniani, R., Crainic, T. G., Gendreau, M., and Rei, W. (2017). The benders decomposition algorithm: A literature review. *European Journal of Operational Research*, 259(3):801–817.

Ropke, S. and Pisinger, D. (2006). An adaptive large neighborhood search heuristic for the pickup and delivery problem with time windows. *Transportation Science*, 40(4):455–472.

Salavati-Khoshghalb, M., Gendreau, M., Jabali, O., and Rei, W. (2019). An exact algorithm to solve the vehicle routing problem with stochastic demands under an optimal restocking policy. *European Journal of Operational Research*, 273(1):175–189.

Secomandi, N. and Margot, F. (2009). Reoptimization approaches for the vehicle-routing problem with stochastic demands. *Operations Research*, 57(1):214–230.

Séguin, R. (1996). *Problèmes stochastiques de tournées de véhicules*. PhD thesis, Université de Montréal, Montréal.

Tillman, F. (1969). The multiple terminal delivery problem with probabilistic demands. *Transportation Science*, 3(3):192–204.

Van Slyke, R. M. and Wets, R. (1969). L-shaped linear programs with applications to optimal control and stochastic programming. *SIAM Journal on Applied Mathematics*, 17(4):638–663.

Wets, R. (1966). Programming under uncertainty: the equivalent convex program. *SIAM Journal on Applied Mathematics*, 14(1):89–105.

Yang, W.-H., Mathur, K., and Ballou, R. (2000). Stochastic vehicle routing problem with restocking. *Transportation Science*, 34(1):99–112.