A comparison of different approaches for the vehicle routing problem with stochastic demands

Ricardo Fukasawa[a], Matheus J. Ota[a,*], Tina Wan[a,1]

[a] University of Waterloo, Waterloo, Ontario, Canada

Abstract

The vehicle routing problem with stochastic demands (VRPSD) is a well studied variant of the classic (deterministic) capacitated vehicle routing problem (CVRP) where the customer demands are given by random variables. Two prominent approaches for solving the VRPSD model it either as a chance-constraint program (CC-VRPSD) or as a two-stage stochastic program (2S-VRPSD). In this paper, we also propose a third combined approach, where we minimize the same objective function as 2S-VRPSD, but over the feasible region of CC-VRPSD. While CC-VRPSD and 2S-VRPSD were extensively studied individually, not much has been investigated regarding the possible advantages/disadvantages of one method against the other. We try to address this gap and conduct theoretical and practical comparisons between these variants of the problem. First, we derive worst-case bounds and show that these bounds can be made arbitrarily close to being tight. We also show sufficient conditions under which chance-constraint feasible routes are feasible for the deterministic counterpart. Next, we implement exact algorithms for solving all the considered approaches and conduct extensive computational experiments to measure their performance. Our findings show that, while the two-stage approach might experience a high failure ratio, in comparison to optimal solutions for the chance-constraint method, the combined approach attains a small failure ratio as well as second-stage cost.

Keywords: Routing, Combinatorial Optimization, Stochastic programming, Integer programming

1. Introduction

Since their introduction by Dantzig & Ramser (1959), vehicle routing problems have been largely investigated, and many different variants have been proposed (Golden et al. (2008)).

*Corresponding author

Email addresses: `rfukasawa@uwaterloo.ca` (Ricardo Fukasawa), `mjota@uwaterloo.ca` (Matheus J. Ota), `zw3720@nyu.edu` (Tina Wan)

[1] Present Address: New York University, New York, NY, USA

In the classic Capacitated Vehicle Routing Problem (CVRP), one is interested in finding a set of routes which collect goods from the customers while respecting the vehicle capacity. Since the CVRP is a generalization of the Travelling Salesperson Problem (TSP), it is NP-hard. Nevertheless, due to its wide applicability, an extensive literature on algorithms for solving the CVRP is available, both from an exact and a heuristic perspective (see Toth & Vigo (2014)).

The CVRP assumes that the decision maker (DM) has access to the precise values of the customer demands, in other words, the demands are *deterministic*. However, in many real-world applications, the demands of the customers are subject to uncertainties, and the DM needs to plan the routes before the uncertainty gets realized. One may attempt to model this situation by replacing stochastic parameters with some deterministic counterparts; however, as shown by Louveaux (1998), this approach can yield poor solutions. Thus, it is important to study the Vehicle Routing Problem with Stochastic Demands (VRPSD), a variant of the CVRP where the demands are random variables. Many approaches have been proposed to solve the VRPSD, here we briefly mention some results concerning exact algorithms (the reader is referred to Oyola et al. (2017, 2018), for a detailed survey on the topic).

When defining an optimization model for the VRPSD, one needs to decide how to handle the possibility that a planned route might *fail*: this is the situation when, after the realization of the random variables corresponding to the demands, the capacity of a vehicle gets exceeded. In this context, two common approaches are to model the VRPSD as a *stochastic program with recourse* (SPR) or as a *chance-constraint program* (CCP). In the SPR approach, the DM defines a recourse policy, which describes the actions that one should take when route failures are observed. Typically, an SPR approach for the VRPSD leads to a *two-stage stochastic program*, where the objective function minimizes the cost of the initial planned routes (first-stage) in addition to the expected cost of executing the recourse policy (second-stage). We call this problem the 2S-VRPSD. On the other hand, the CCP approach ignores the recourse cost, but it sets an upper bound of $\epsilon \in (0, 1)$ on the probability that a route observes a failure. We use CC-VRPSD to denote the chance-constraint version of the VRPSD.

The first exact solving method for the 2S-VRPSD was based on cutting planes and was proposed by Gendreau et al. (1995). The authors designed a branch-and-cut algorithm based on the *integer L-shaped method*. Let us denote by $Q(x)$ the recourse cost associated with x, where x is a vector encoding a feasible solution for the VRPSD. In the integer L-shaped method, we have a mathematical formulation where a non-negative real variable θ lower bounds the second-stage cost. Then, whenever an integer solution $\bar{x}$ is found, we compute the second-stage cost $Q(\bar{x})$ and add an *optimality cut*, which enforces that $\theta = Q(\bar{x})$ when $x = \bar{x}$. Usually the added optimality cut does not provide good lower bounds for θ at solutions distinct from $\bar{x}$. To address this issue, Laporte et al. (2002) and Jabali et al. (2014) proposed the addition of optimality cuts based on *lower bounding functionals* (LBF's), which may provide tighter approximations for the recourse cost function $Q(x)$ by adding a smaller number of optimality cuts to the model. Indeed, their computational experiments show that LBF's improve considerably the performance of branch-and-cut algorithms based on the integer L-shaped method. Column generation approaches were also proposed for the 2S-VRPSD (see Christiansen & Lysgaard (2007); Gauvin et al. (2014)), where the second-stage cost is incorporated in the objective function coefficient of a column. We also remark that all of the mentioned works use the so called *classical* or *detour-to-depot* recourse policy. That is, when executing a route, a vehicle replenishes its capacity — by executing back-and-forth trips to the depot — only when the demand of a customer exceeds the vehicle's residual capacity. Naturally, more sophisticated recourse policies have been proposed in the literature. For example, Yee & Golden (1980) proposed an *optimal restocking policy* and Florio et al. (2020) designed a branch-cut-and-price algorithm for the 2S-VRPSD under this policy. More recently, Salavati-Khoshghalb et al. (2019b,a) considered *rule-based* recourse policies.

In comparison to the two-stage approach, the literature for the chance-constraint VRPSD is considerably smaller. Stewart & Golden (1983) were the first to propose an exact algorithm for CC-VRPSD. Under very specific conditions on the probability distributions of the demands, they show that CC-VRPSD can be reduced to the deterministic CVRP by modifying the vehicle capacity. Later, Laporte et al. (1989) designed a branch-and-cut algorithm for CC-VRPSD when the customer demands follow independent normal distributions. Their algorithm models the chance-constraints by modifying the right-hand side coefficient of the classic rounded capacity inequalities (RCI) of Laporte & Nobert (1983). All of the algorithms mentioned so far assume that the demands are independent random variables; however, in real-world applications, often the customers demands are correlated. In fact, in many settings, the decision maker only has access to demand *scenarios* — which tracks the clients demands in the past — and no assumption can be made concerning the independence of the demands random variables. Dinh et al. (2018) make an important improvement in this sense. They extend the approach of Laporte et al. (1989) of modifying the RCI's for a much more general case, in particular, their method does not assume that the customer demands are independent random variables. The only requirement needed by their approach is that the DM can compute a quantile of the accumulated demand of any subset of customers. This implies that their algorithm also works for the previously mentioned scenario model.

Finally, in this paper, we introduce a third model for handling the stochasticity, which we call the *two-stage chance-constraint vehicle routing problem with stochastic demands* (2SCC-VRPSD). In this approach, we aim to minimize the total cost (accounting for the recourse costs) over the same feasible region as in CC-VRPSD. For short, we often refer to this model simply as *combined*.

Although the recent literature shows advances in either the two-stage or the chance-constraint approach individually, less attention have been given to comparing the quality of the solutions found by both methods. In fact, to the best of our knowledge, the only work that presents an explicit comparison of the methods is an appendix in the paper of Dinh et al. (2018), where they conducted computational experiments in a very limited set of instances. In this work, we conduct theoretical and computational comparisons of the investigated models. On the theoretical side, we first consider solutions over the same feasible region, and prove worst-case bounds on the cost of a solution which ignores the second-stage cost, relative to approaches that take such costs into account. In other words, we derive bounds comparing the total cost of optimal solutions for the deterministic CVRP and 2S-VRPSD, as well as bounds for the total cost of CC-VRPSD against 2SCC-VRPSD. We also show instances where these bounds can be made arbitrarily close to being tight. Furthermore, we derive sufficient conditions under which chance-constraint feasible routes are feasible for CVRP (or 2S-VRPSD). In the second part of the paper, we conduct extensive computational experiments to investigate the quality of the solutions for all the three methods. Our findings show that even when the instances are randomly generated, with correlations or not, the two-stage approach might experience a high failure ratio. On the other hand, the combined approach attains a small failure ratio as well as second-stage cost.

2. Preliminaries

As usual, if t is a positive integer, $[t] := \{1, \ldots, t\}$; moreover, when $t \leq 0$, $[t]$ indicates the empty set. If g is a vector and i is one of its coordinates, we may write interchangeably g_i or $g(i)$. For any function (resp. vector) f, where H is a subset of its domain (resp. subset of its coordinates), we use $f(H)$ as a short-hand for $\sum_{i \in H} f(i)$. Moreover, when G' is a graph or a digraph, we write $f(G')$ meaning $f(V(G'))$. For ease of presentation, we sometimes abbreviate an edge $\{i, j\}$ or an arc (i, j) simply to ij. Since the numbers in our input data are rational numbers, we may assume without loss of generality that they are integers.

Let $G = (V, E)$ be a graph, where $V = \{0\} \cup V_+$. Vertex 0 represents the depot and the set V_+ denotes the set of customers. For each $i \in V$, the notation $\delta(i)$ indicates the set of edges in G incident to i. For each edge $e \in E$, c_e is a positive integer indicating the cost of edge e. Moreover, $k \in \mathbb{Z}_+$ and $C \in \mathbb{Z}_+$ denote the desired number of routes and the vehicle capacity, respectively.

Let S be a finite set of scenarios. For each $s \in S$, $d^s \in \mathbb{Z}_+^{V_+}$ is a vector where each entry d_i^s indicates the demand of customer $i \in V_+$ in scenario s. The scenarios also have associated probabilities, to this end, we define p_s as the realization probability of a scenario $s \in S$. We denote by d the random vector with $\mathsf{P}(d = d^s) = p_s$, for all $s \in S$. Additionally, $\bar{d} \in \mathbb{Z}_+^{V_+}$ represents the expected demands of the customers, that is, $\bar{d} = \mathsf{E}[d]$.

In the vehicle routing problem literature, one naturally needs to speak of the routes taken by the vehicles. Additionally, when considering a recourse policy, one also needs to take into account the two different directions associated with a route. For preciseness, we now formalize these concepts. We say that a *route* $R \subseteq G$ is a simple undirected cycle which contains the depot, that is, $V(R) = \{0, v_1, v_2, \dots, v_\ell\}$ and $E(R) = \{\{0, v_1\}, \{v_1, v_2\}, \dots, \{v_\ell, 0\}\}$. For ease of presentation, we write $R = (v_1, \dots, v_\ell)$. Furthermore, we use $c(R)$ to refer to the sum $\sum_{e \in E(R)} c_e$ and $d(R)$ to refer to $d(V(R) \setminus \{0\})$. Each route $R = (v_1, \dots, v_\ell)$ has two associated digraphs $\vec{R}$ and $\overleftarrow{R}$, which we call *directed routes*. Both $\vec{R}$ and $\overleftarrow{R}$ have the same vertex set as R, but the arcs are in opposite directions, that is, $A(\vec{R}) = \{(0, v_1) \dots, (v_\ell, 0)\}$ and $A(\overleftarrow{R}) = \{(0, v_\ell) \dots, (v_1, 0)\}$. Similarly to the routes, we write $\vec{R} = (v_1, \dots, v_\ell)$ and $\overleftarrow{R} = (v_\ell, \dots, v_1)$.

In the (deterministic) capacitated vehicle routing problem (CVRP), we aim to find k routes $R_1, \dots, R_k$ such that

(i) $\{V(R_i) \setminus \{0\}\}_{i \in [k]}$ forms a partition of V_+;

(ii) $\bar{d}(R_i) \leq C$, for all $i \in [k]$; and

(iii) $\sum_{i \in [k]} c(R_i)$ is minimum.

Notice that we are using $\bar{d}_i$ as the demand of customer $i \in V_+$. If $\{R_i\}_{i \in [k]}$ satisfy (i) and (ii) we say that $\{R_i\}_{i \in [k]}$ is a *deterministic feasible solution*. Similarly, if a route R satisfies (ii) we say that R is a *deterministically feasible route*. This is not a standard nomenclature, but it will be useful for our purposes.

Using binary variables x_e, for each edge $e \in E$, we can define a set $X \subseteq \mathbb{R}^E$, which represents the set of feasible solutions for the CVRP. One typical way to formulate the CVRP uses the rounded capacity inequalities (RCI's) proposed by Laporte & Nobert (1983):

$$X = \left\{ x \in \{0,1,2\}^E : \begin{array}{ll} x(\delta(0)) = 2k, & \\ x(\delta(i)) = 2, & \forall i \in V_+, \\ x(\delta(U)) \geq 2k(U) & \forall U \subseteq V_+, U \neq \emptyset, \\ x_e \in \{0,1\}, & \forall e \in E \setminus \delta(0), \end{array} \right\} \tag{1}$$

where $k(U) := \lceil \bar{d}(U)/C \rceil$. It follows that a formulation for the CVRP is given by $\min\{c^T x : x \in X\}$. We say that $x \in \mathbb{R}^E$ is *deterministically feasible* if $x \in X$.

Since some of the problems studied will use routes that are not deterministically feasible, it is worth defining $\tilde{X}$ to be obtained from X by replacing $k(U)$ with 1, for all $\emptyset \neq U \subseteq V_+$. Vectors in $\bar{x} \in \tilde{X} \supseteq X$ correspond to solutions without subcycles; therefore, with each $\bar{x} \in \tilde{X}$, we can associate a set of routes $R(\bar{x}) = \{R_1(\bar{x}), \dots, R_k(\bar{x})\}$ in the obvious way. Observe that if $x_1 \in X$ and $x_2 \in \tilde{X}$, then $R(x_1)$ satisfy (i) and (ii), while $R(x_2)$ is only guaranteed to satisfy (i).

In what follows, we assume that when a failure is observed, the recourse policy is to do a back-and-forth trip between the depot and the failed customer, i.e, we assume the detour-to-depot recourse policy. In this context, Laporte et al. (2002) show that for any directed route $\vec{R} = (v_1, \ldots, v_\ell)$, the expected recourse cost $Q(\vec{R})$ can be computed with the formula

$$Q(\vec{R}) = 2 \sum_{j \in [\ell]} \sum_{t=1}^{\infty} \mathsf{P}\left(\sum_{i \in [j-1]} d(v_i) \le tC < \sum_{i \in [j]} d(v_i) \right) c_{0 v_j}.$$

The recourse cost of an (undirected) route R is then given by $Q(R) = \min\{Q(\vec{R}), Q(\overleftarrow{R})\}$. In the two-stage vehicle routing problem with stochastic demands (2S-VRPSD) we want to find routes $R_1, \ldots, R_k$ that satisfies (i) and (ii), but minimizes $\sum_{i \in [k]} (c(R_i) + Q(R_i))$. In other words, we aim to solve the following optimization problem: $\min\{c^T x + Q(x) : x \in X\}$, where $Q(x) = \sum_{R \in \mathcal{R}(x)} Q(R)$. For any vector $\bar{x} \in \tilde{X}$, we say that $c^T \bar{x}$ is its *first-stage cost*, $Q(\bar{x})$ is its *second-stage cost* or *recourse cost*, and the sum $c^T \bar{x} + Q(\bar{x})$ is its *total cost*.

Sometimes we will also need to refer to the recourse costs associated with a specific scenario $s \in S$. Let $\vec{R} = (v_1, \ldots, v_\ell)$ be a directed route. For each $s \in S$, $Q(\vec{R}, s)$ is the recourse cost of $\vec{R}$ w.r.t. scenario s and is given by

$$Q(\vec{R}, s) = 2 \sum_{j \in [\ell]} \sum_{t=1}^{\infty} \mathsf{I}\left(\sum_{i \in [j-1]} d^s(v_i) \le tC < \sum_{i \in [j]} d^s(v_i) \right) c_{0 v_j}.$$

The recourse cost of $\bar{x} \in \tilde{X}$ with respect to s can then be obtained with the formula $Q(\bar{x}, s) = \sum_{R \in \mathcal{R}(\bar{x})} \min\{Q(\vec{R}, s), Q(\overleftarrow{R}, s)\}$. From now on, we say that a route R *fails* in scenario $s \in S$, if $d^s(R) > C$.

In the chance-constraint vehicle routing problem with stochastic demands (CC-VRPSD) we handle the case where the demands are stochastic by enforcing that a route will violate the capacity constraint with a low probability. Fix a tolerance constant $\epsilon \in (0, 1)$, in the CC-VRPSD we seek k routes $R_1, \ldots, R_k$ satisfying (i) and (iii), but we replace (ii) with the condition

(ii') $\mathsf{P}(d(R_i) > C) \le \epsilon$, for all $i \in [k]$.

Similar to the deterministic case, if $\{R_i\}_{i \in [k]}$ satisfy (i) and (ii'), we say that $\{R_i\}_{i \in [k]}$ is a *chance-constraint feasible solution*; and if a route R satisfies (ii') we say that R is *chance-constraint feasible route*. Let $X_{\mathrm{CC}}(\epsilon) \subseteq \mathsf{R}^E$ be the feasible region of CC-VRPSD (when ϵ is clear from the context, we write simply X_{CC}), that is, each $x \in X_{\mathrm{CC}}$ corresponds to a feasible chance-constraint solution. Dinh et al. (2018) prove that X_{CC} can be modeled by replacing the $k(U)$ in the formulation for X by a value $\rho(U)$, for all $\emptyset \ne U \subseteq V_+$. We call the resulting inequalities the *chance-constraint rounded capacity inequalities* (CC-RCI's). Finally, the two-stage chance-constraint vehicle routing problem with stochastic demands (2SCC-VRPSD) is defined as $\min\{c^T x + \mathcal{Q}(x) : x \in X_{\mathrm{CC}}\}$. In other words, we seek k routes $R_1, \ldots, R_k$ that satisfies (i), (ii') and minimizes $\sum_{i\in[k]}(c(R_i) + \mathcal{Q}(R_i))$.

We now discuss how we incorporate the second-stage cost into a formulation for VRPSD. Recall that the lower bounding functionals shown by Laporte et al. (2002) and Jabali et al. (2014) assume that the probability distributions of the customer demands are independent. Since we do not make such an assumption, we approximate the recourse cost using the traditional *integer L-shaped method*. For completeness, we briefly describe its details. The reader is referred to Gendreau et al. (1995) for a better coverage of the topic.

Let $y_s \geq 0$, for all $s \in S$, be a continuous variable indicating the recourse cost associated with scenario s. Gendreau et al. (1995) proved the following.

Proposition 1. *Let $\bar{x} \in X$, then, for every scenario $s \in S$, the following is a valid optimality cut for the 2S-VRPSD*

$$y_s \geq \mathcal{Q}(\bar{x}, s)(x(\bar{E} \setminus \delta(0)) - |V| + k + 2), \tag{2}$$

where $\bar{E} = \{e \in E : \bar{x}_e > 0\}$.

Proof. Let $R_i = (v^i_1, \ldots, v^i_{\ell_i})$, for all $i \in [k]$, be the routes associated with the solution $\bar{x}$. Consider the simple path P made of the vertices in V_+ in the following order

$$P = (v^1_1, \ldots, v^1_{\ell_1}, v^2_1, \ldots, v^2_{\ell_2}, \ldots, v^k_1, \ldots, v^k_{\ell_k}).$$

Then, $\bar{x}(\bar{E}\setminus\delta(0)) = |\bar{E}\setminus\delta(0)| = |E(P)| - (k-1) = |V| - k - 1$, and it follows that $\mathcal{Q}(\bar{x}, s)(\bar{x}(\bar{E}\setminus\delta(0)) - |V| + k + 2) = \mathcal{Q}(\bar{x}, s)$. Now take $x' \in X$ with $x' \neq \bar{x}$, then $x'(\bar{E}\setminus\delta(0)) < |\bar{E}\setminus\delta(0)|$, and therefore, $x'(\bar{E}\setminus\delta(0)) - |V| - k - 2$. □

Using Proposition 1 we have the following stochastic mixed integer program formulation for the 2S-VRPSD.

$$\begin{aligned}
\text{(ILS)} \quad \min \quad & \sum_{e\in E} c(e)x_e + \sum_{s\in S} p(s)y_s && \\
\text{s.t.} \quad & x \in X, && (3)\\
& y_s \geq \mathcal{Q}(\bar{x}, s)(x(\bar{E}\setminus\delta(0)) - |V| + k + 2), \qquad \forall s \in S, \bar{x} \in X, && (4)\\
& y_s \geq 0, \qquad \forall s \in S. && (5)
\end{aligned}$$

Since there may be exponentially many inequalities of type (4), we treat these with a cutting-plane approach. Notice that it follows from the proof of Proposition 1 that (4) is tight only at $x = \bar{x}$, as otherwise it is at least as weak as the inequality $y_s \geq 0$. In order to solve 2SCC-VRPSD it suffices to replace $\mathcal{X}$ by $\mathcal{X}_{\text{CC}}$ in (ILS); in practice, this means that the cutting-plane algorithm separates CC-RCI's instead of RCI's.

3. Theoretical bounds

In this section, we assume that G and c satisfy the triangle inequality. Moreover, we denote by $\bar{x}$ and $\bar{x}_{2\text{S}}$ optimal solutions for CVRP and 2S-VRPSD. Similarly, $\bar{x}_{\text{CC}}$ and $\bar{x}_{2\text{SCC}}$ denote optimal solutions for CC-VRPSD and 2SCC-VRPSD.

Our first result is a bound that relates how costly can a solution be if we do not consider the second-stage cost. This bound depends on how many failures happen at optimal solutions to the problem of minimizing just the first-stage cost. Although simple, this bound will allow us to derive other results later on.

Theorem 2. *Let $t \geq 1$ and suppose that each route R in $\mathcal{R}(\bar{x})$ fails at most t times, then*

$$c^T\bar{x} + Q(\bar{x}) \leq (t+1) \cdot (c^T\bar{x}_{2\text{S}} + Q(\bar{x}_{2\text{S}})).$$

Proof. Since $\bar{x}$ is optimal for the deterministic CVRP, $c^T\bar{x} \leq c^T\bar{x}_{2\text{S}}$. In addition, by our assumption on the maximum number of failures of a route, it follows from the triangle inequality that $Q(R) \leq t \cdot c(R)$, for each $R \in \mathcal{R}(\bar{x})$. Therefore $Q(\bar{x}) \leq t \cdot c^T\bar{x}$ and we have that

$$c^T\bar{x} + Q(\bar{x}) \leq (t+1) \cdot c^T\bar{x} \leq (t+1) \cdot c^T\bar{x}_{2\text{S}} \leq (t+1) \cdot (c^T\bar{x}_{2\text{S}} + Q(\bar{x}_{2\text{S}})).$$

□

Some remarks should be made regarding Theorem 2. First, one can easily check that the theorem also holds for $\mathcal{X}_{\text{CC}}$, that is, it is valid if we replace $\bar{x}$ and $\bar{x}_{2\text{S}}$ by $\bar{x}_{\text{CC}}$ and $\bar{x}_{2\text{SCC}}$. Second, our experiments (detailed on Section 4) show that, in practice, many optimal solutions of CVRP or 2S-VRPSD satisfy the property that all of its routes fail at most once; and in such settings, Theorem 2 states that the total cost of an optimal deterministic solution is at most twice the total cost of a two-stage solution. Proposition 3 below attempts to formalize this observation for the case when the accumulated demand of a route follows a normal distribution with a variance that is not too large. Although Proposition 3 seems to assume too specific assumptions, Corollary 4 shows that it encompasses the case where the demands are independent and normally distributed with variances that are not too large (which is a case considered by a large number of works in the literature).

Proposition 3. *Let $R_1, \ldots, R_k$ be all the routes in $\mathcal{R}(\bar{x})$. For each $i \in [k]$, assume $d(R_i)$ follows a normal distribution $N(\mu_i, \sigma_i)$ and $\sigma_i \leq \mu_i/2$. Then, with probability at least $1 - 0.025k$, each route in $\mathcal{R}(\bar{x})$ fails at most once.*

Proof. Fix $i \in [k]$ and note that $\mu_i \leq C$, since $\bar{x} \in X$. Hence, $\mu_i + 2\sigma_i \leq 2\mu_i \leq 2C$, and we have that $\mathsf{P}(d(R_i) \geq 2C) \leq \mathsf{P}(d(R_i) \geq \mu_i + 2\sigma_i) \leq 0.025$, where the last inequality follows from the fact that $\mathsf{P}(|d(R_i) - \mu_i| \geq \mu_i + 2\sigma_i) \leq 5\%$. The result then follows from the union bound. □

Corollary 4. *Suppose $\{d_i\}_{i \in V_+}$ are independent and for each $i \in V_+$, d_i follows a normal distribution $N(\mu_i, \sigma_i^2)$, with $\sigma_i \leq \mu_i/2$. Then, with probability at least $1 - 0.025k$, each route in $\mathcal{R}(\bar{x})$ fails at most once.*

Proof. Fix $R \in \mathcal{R}(\bar{x})$. Let $\mu = \sum_{i \in V(R)} \mu_i$ and $\sigma^2 = \sum_{i \in V(R)\setminus\{0\}} \sigma_i^2$. Since the square root function is subadditive for non-negative arguments, $\sigma \leq \sum_{i \in V(R)\setminus\{0\}} \sigma_i$. As $d(R)$ is obtained through the sum of normal distributions, it follows a normal distribution $N(\mu, \sigma^2)$. Moreover, $2\sigma \leq 2\sum_{i \in V(R)\setminus\{0\}} \sigma_i \leq \mu$. We are now done by Proposition 3. □

As we previously mentioned, Theorem 2 also holds for CC-VRPSD, however, in this case, one can use the tolerance parameter ϵ to derive better bounds.

Theorem 5. *Let $t \geq 1$ and suppose that each route R in $\mathcal{R}(\bar{x}_{\mathrm{CC}})$ fails at most t times, then*

$$c^T \bar{x}_{\mathrm{CC}} + Q(\bar{x}_{\mathrm{CC}}) \leq (1 + t\epsilon) \cdot (c^T \bar{x}_{\mathrm{2SCC}} + Q(\bar{x}_{\mathrm{2SCC}})).$$

Proof. Since $\bar{x}_{\mathrm{CC}} \in X_{\mathrm{CC}}(\epsilon)$, it follows from the theorem assumption and the triangle inequality that $Q(R) \leq t\epsilon \cdot c(R)$, for all $R \in \mathcal{R}(\bar{x}_{\mathrm{CC}})$. Thus, $Q(\bar{x}_{\mathrm{CC}}) \leq t\epsilon \cdot c^T \bar{x}_{\mathrm{CC}}$. Overall, we conclude that

$$c^T \bar{x}_{\mathrm{CC}} + Q(\bar{x}_{\mathrm{CC}}) \leq (1 + t\epsilon) \cdot c^T \bar{x}_{\mathrm{CC}} \leq (1 + t\epsilon) \cdot (c^T \bar{x}_{\mathrm{2SCC}} + Q(\bar{x}_{\mathrm{2S}})).$$

□

Corollary 6. *The inequality $c^T \bar{x}_{\mathrm{CC}} + Q(\bar{x}_{\mathrm{CC}}) \leq (1 + t\epsilon) \cdot (c^T \bar{x}_{\mathrm{2SCC}} + Q(\bar{x}_{\mathrm{2SCC}}))$ holds with probability at least $1 - k\epsilon$.*

Proof. Fix a route $R \in \mathcal{R}(\bar{x}_{\mathrm{CC}})$. Since $\bar{x}_{\mathrm{CC}} \in X_{\mathrm{CC}}$, the probability that R has at least two failures is at most ϵ. Thus, the probability that there is a route in $\mathcal{R}(\bar{x}_{\mathrm{CC}})$ with at least two failures is at most $k\epsilon$. Hence, the result follows from Theorem 5. □

Observe that Theorems 2 and 5 remain valid if we replace $Q(x)$ by the cost of any recourse policy on x, as long as the cost of t failures is at most $t \cdot c(R_i)$ for any route $R_i \in \mathcal{R}(x)$. Moreover, at first sight, the analysis seems loose, since the inequalities are fairly simple. However, under the detour-to-depot policy and with $t = 1$, we show that the upper bound can be made arbitrarily close to being tight. In order to do so, we first define a specific type of graph.

Definition 7. A *circle graph* is given by a tuple $\mathsf{C} = (G^\circ, c^\circ, n, r, \sigma)$, where $n \in \mathbb{Z}_+$, $r \in \mathbb{R}_+$, $\sigma \in \mathbb{R}_+$, $G^\circ = (V^\circ, E^\circ)$ is an euclidean graph drawn in the plane and $c^\circ \in \mathbb{R}_+^{E^\circ}$. The set of vertices is $V^\circ = \{v_0\} \cup V_+^\circ$, where v_0 is the depot and $V_+^\circ = \{v_i\}_{i \in [n^2]}$ are the customers. The vertex v_0 is positioned at the origin, while the customers $v_1, \ldots, v_{n^2}$ are positioned along a circumference of radius r centered at the origin. In addition, for each $i \in [n^2 - 1]$, the distance between v_i and v_{i+1} is σ. Since the graph G° is euclidean, it is complete and the edge costs c° is determined by the vertex positions (see Figure 1).

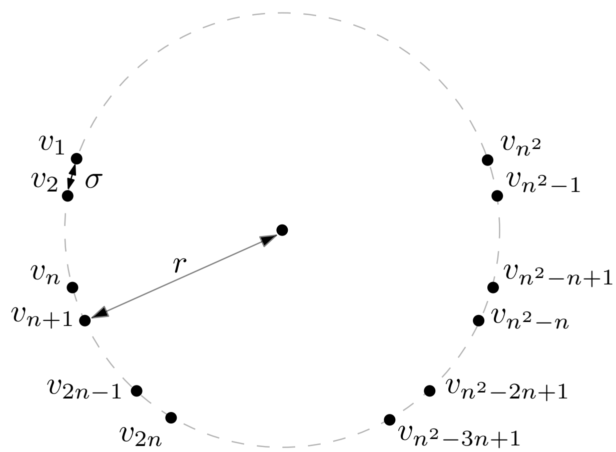

Figure 1: Example of a circle graph.

In what follows, whenever we consider a circle graph, we assume that its parameters are chosen in a way that the unique closest vertex to v_1 (resp. v_{n^2}) is v_2 (resp. v_{n^2-1}). In the next results, for a given instance I of VRPSD, we write $X(I)$ to refer to the formulation of X with respect to instance I.

Theorem 8. *Let $\Delta > 0$ be an arbitrarily small constant. Then there exists an instance I with costs c such that each route $R \in \mathcal{R}(\bar{x})$ fails at most once and $c^T\bar{x} + Q(\bar{x}) \geq (2 - \Delta)(c^T\bar{x}_{2\mathrm{S}} + Q(\bar{x}_{2\mathrm{S}}))$, where $\bar{x}$ and $\bar{x}_{2\mathrm{S}}$ are minimizers of $c^T x$ and $c^T x + Q(x)$ over $X(I)$.*

Proof. Let $\mathsf{C} = (G^\circ, c^\circ, n, r, \sigma)$ be a circle graph, where $r >> \sigma$ and n is large. We set $k = n$ and we claim that a route visiting n customers has minimum cost if and only if these customers are visited sequentially. To see this, let $i_1, \ldots i_n$ be such that $1 \leq i_1 < i_2 < \ldots < i_n \leq n^2$. Let R be a route in G° such that $V(R) = \{v_{i_1}, \ldots, v_{i_n}\}$. The cost of R is given by $c^\circ(E(R)) = 2r + \sum_{j \in [n-1]} c^\circ((v_{i_j}, v_{i_{j+1}})) \geq 2r + (n-1)\sigma$, where the inequality follows from the fact that every edge in G° has cost at least σ. Moreover, by the ordering of $i_1, \ldots, i_n$, equality holds if and only if $i_{j+1} = i_j + 1$, for all $j \in [n-1]$. Therefore, $c^\circ(E(R)) \geq 2r + (n-1)\sigma$ and the equality is reached if and only if $i_{j+1} = i_j + 1$, for all $j \in [n-1]$.

Next, we set the capacity to n^2; moreover, we define the set S containing n scenarios $s_1, \ldots, s_n$. For each $i \in [n]$, we set d^{s_i} as following.

$$d^{s_i}(v) = \begin{cases} 1, & \text{if } v = v_{(i-1)n+j}, \text{ for some } j \in [n], \\ n+1, & \text{otherwise.} \end{cases} \tag{6}$$

Each scenario s_i has the same probability $p_{s_i} = \frac{1}{n}$.

Notice that the expected demand of each client is n. Thus, since $\bar{x}$ is feasible for X, each route visits precisely n customers. By the previous reasoning, the first-stage cost of $\bar{x}$ is $2nr + n(n-1)\sigma$ and $R(\bar{x}) = \{R_i\}_{i\in[n]}$, with $R_i = (v_{(i-1)n+1}, \ldots, v_{(i-1)n+n})$, for all $i \in [n]$. By the way we set the demands, for each $i \in [n]$, R_i fails once in exactly $n-1$ scenarios, and in each such failure it yields a recourse cost of $2r$. Hence, $Q(\bar{x}) = 2nr - 2r$ and $(c^\circ)^T \bar{x} + Q(\bar{x}) = 2nr + 2r(n-1) = 4nr + n(n-1)\sigma - 2r$.

Now let x' be such that $R(x') = \{R'_i\}_{i\in[n]}$ with $R'_i = (v_i, v_{n+i}, v_{2n+i}, \ldots, v_{n^2-n+i})$, for all $i \in [n]$. Notice that no recourse action is taken by x', and thus, by the triangle inequality, $(c^\circ)^T x' + Q(x') \leq 2nr + n^2(n-1)\sigma$. Assume we choose n large enough so that $\Delta n \geq 1$ and σ so that

$$\begin{aligned} & \sigma \leq \frac{2r(\Delta n - 1)}{n(n-1)(2n-1-\Delta n)}, \\ \iff\ & n(n-1)(2n-1-\Delta n)\sigma \leq 2\Delta nr - 2r, \\ \iff\ & n(n-1)(\Delta n + 1 - 2n)\sigma \geq 2r - 2\Delta nr, \\ \iff\ & n(n-1)\sigma - 2r \geq (2-\Delta)n^2(n-1)\sigma - 2\Delta nr. \end{aligned}$$

It follows then that

$$\begin{aligned} (2-\Delta)((c^\circ)^T \bar{x}_{2\mathrm{S}} + Q(\bar{x}_{2\mathrm{S}})) &\leq 4nr - 2\Delta nr + (2-\Delta)n^2(n-1)\sigma, \\ &\leq 4nr + n(n-1)\sigma - 2r, \\ &= (c^\circ)^T \bar{x} + Q(\bar{x}). \end{aligned}$$

□

Theorem 9. *Let $0 < \Delta < 1$ be an arbitrarily small constant. Then there exists an instance I with costs c such that each route $R \in R(\bar{x}_{\mathrm{CC}})$ fails at most once and $c^T \bar{x}_{\mathrm{CC}} + Q(\bar{x}_{\mathrm{CC}}) \geq (1+\epsilon-\Delta)(c^T \bar{x}_{2\mathrm{SCC}} + Q(\bar{x}_{2\mathrm{SCC}}))$, where ϵ is the chance-constraint parameter and $\bar{x}_{\mathrm{CC}}, \bar{x}_{2\mathrm{SCC}}$ are minimizers of $c^T x$ and $c^T x + Q(x)$ over $X_{\mathrm{CC}}(I)$.*

Proof. Let $n \in \mathbb{Z}_+$ be such that ϵn is integer and n is large. Construct sets of scenarios S_1 and S_2 such that $S = S_1 \cup S_2$, $|S_1| = (1-\epsilon)n$ and $|S_2| = \epsilon n$. Similarly to Theorem 8, let $C = (G^\circ, c^\circ, n, r, \sigma)$ be a circle graph, where $r \gg \sigma$, the vehicle capacity is set to n^2

and $k = n$. For each scenario $s_1 \in S_1$, we set $d^{s_1}(v) = n$, for all $v \in V_+$. On the other hand, for each scenario $s_2 \in S_2$, we set

$$d^{s_2}(v) = \begin{cases} n+1, & \text{if } v = v_j, \text{ for some } j \in [n^2 - n], \\ 1, & \text{otherwise.} \end{cases} \tag{7}$$

By the way we set the demands for scenarios in S_1 we know that $\bar{x}_{\text{CC}}$ is feasible for $s_1 \in S_1$ only if each of its routes has precisely n customers. Now take $s_2 \in S_2$ and assume that $\bar{x}_{\text{CC}}$ is feasible for s_2. Since $\sum_{v \in V} d^{s_2}(v) = n^3$, by the pigeonhole principle, each route in $\mathcal{R}(\bar{x}_{\text{CC}})$ has accumulated demand of precisely n^2. Moreover, for each route $R \in \mathcal{R}(\bar{x}_{\text{CC}})$, it holds that $|V(R) \cap \{v_1, \dots, v_{n^2-n}\}| = n - 1$. This implies that if $\bar{x}_{\text{CC}}$ is feasible for a scenario in S_2 then each of its routes has exactly n customers.

Next, we claim that $\mathcal{R}(\bar{x}_{\text{CC}}) = \{(v_1, \dots, v_n), (v_{n+1}, \dots, v_{2n}), \dots, (v_{n^2-n+1}, \dots, v_{n^2})\}$. Indeed, recall from Theorem 8 that a route that visits n customers have cost at least $2r + (n-1)\sigma$, and the lower bound is attained if and only if the customers are visited sequentially in the route. Hence,

$$(c^\circ)^T \bar{x}_{\text{CC}} + Q(\bar{x}_{\text{CC}}) = 2nr + n(n-1)\sigma + \frac{2r|S_2|(n-1)}{|S|} = 2nr + n(n-1)\sigma + 2r\epsilon(n-1). \tag{8}$$

Consider now x'_{CC} such that

$$\mathcal{R}(x'_{\text{CC}}) = \{(v_1, \dots, v_{n-1}, v_{n^2-n+1}), \dots, ((v_{(n-1)^2}, \dots, v_{n(n-1)}, v_{n^2})\},$$

and notice that x'_{CC} has no recourse cost. Thus, $(c^\circ)^T x'_{\text{CC}} \leq 2nr + n(n^2 - n)\sigma$. Now choose n such that $\epsilon/n < \Delta$ and σ such that $(1 + \epsilon - \Delta)n(n^2 - n)\sigma \leq 2nr\Delta - 2r\epsilon$, then

$$\begin{aligned} (c^\circ)^T \bar{x}_{\text{CC}} + Q(\bar{x}_{\text{CC}}) &\geq 2nr + 2r\epsilon(n-1) && \text{(by (8))} \\ &\geq 2nr + 2r\epsilon n + (1+\epsilon-\Delta)n(n^2-n)\sigma - 2nr\Delta && \text{(by the choice of } \sigma) \\ &= (1+\epsilon-\Delta)(2nr + n(n^2-n)\sigma) \\ &\geq (1+\epsilon-\Delta)(c^\circ)^T x'_{\text{CC}}. \end{aligned}$$

$\square$

The previous comparisons are all between solutions which belongs to the same feasible region: either $\mathcal{X}$ or $\mathcal{X}_{\text{CC}}$. Consider now that we want to compare $\bar{x}_{2\text{S}} \in \mathcal{X}$ with $\bar{x}_{\text{CC}} \in \mathcal{X}_{\text{CC}}$. If it is the case that $\bar{x}_{\text{CC}}$ lies in $\mathcal{X}$, then we know that the total cost of $\bar{x}_{2\text{S}}$ is a lower bound for the total cost of $\bar{x}_{\text{CC}}$, i.e., $c^T \bar{x}_{2\text{S}} + Q(\bar{x}_{2\text{S}}) \leq c^T \bar{x}_{\text{CC}} + Q(\bar{x}_{\text{CC}})$. Our next goal is to investigate situations where this might hold. Let R be a chance-constraint feasible route and suppose that $d(R)$ is upper bounded by $2C$. In this case, one may expect that R is deterministically feasible if ϵ is small, however, this is not true in general. Indeed, suppose that R is a route such that $d(R) = C$ with probability $1 - \epsilon$ and $d(R) > C$ with probability ϵ. Then, for any $\epsilon > 0$, we have that R is chance-constraint feasible, but $\bar{d}(R) > C$. The next result shows that, although we have no hope of guaranteeing that a chance-constraint feasible route R is deterministically feasible in general, we can guarantee (under some conditions on $\mathsf{E}[d(R)|d(R) > C]$) that R is deterministically feasible for a slightly larger capacity than C. Similarly, if R is chance-constraint feasible for a capacity slightly smaller than C, then R is also deterministically feasible for C (set $p = 1 - \epsilon$ in Theorem 10 below). In what follows, for a given random variable Z, we use $Q_p(Z) := \inf\{\alpha : \mathsf{P}(Z \leq \alpha) \geq p\}$ to denote its p-th quantile.

Theorem 10. *Let R be a chance-constraint feasible route and suppose $\mathsf{E}[d(R)|d(R) > C] \leq C(1+q)$, where $q \geq 0$. Then, the following hold.*

(a) $\bar{d}(R) \leq C(1+\epsilon q)$;

(b) if there is $p \in (0,1)$ such that $Q_p(d(R)) \geq C\left(1 - \frac{q}{p}\right)$, then $\bar{d}(R) \leq C$.

Proof. To ease notation, we define $\bar{d} := \bar{d}(R)$. Item (a) follows from the law of total expectation,

$$\begin{aligned}
\bar{d} &= \mathsf{E}[d(R)|d(R) \leq C]\,\mathsf{P}(d(R) \leq C) + \mathsf{E}[d(R)|d(R) > C]\,\mathsf{P}(d(R) > C) \\
&\leq \mathsf{E}[d(R)|d(R) \leq C]\,(1-\epsilon) + \mathsf{E}[d(R)|d(R) > C]\,\epsilon \\
&\leq C(1-\epsilon) + C(1+q)\epsilon \\
&= C(1+\epsilon q).
\end{aligned}$$

For (b), we start by writing $\bar{d}$ as a weighted sum (here we assume $d(R)$ is a discrete random variable, but the argument also holds for continuous random variables),

$$\begin{aligned}
& \bar{d} = \sum_{t=0}^{\bar{d}} t \cdot \mathsf{P}(d(R)=t) + \sum_{t=\bar{d}+1}^{\infty} t \cdot \mathsf{P}(d(R)=t) \\
\iff\ & 0 = \sum_{t=0}^{\bar{d}} (t-\bar{d}) \cdot \mathsf{P}(d(R)=t) + \sum_{t=\bar{d}+1}^{\infty} (t-\bar{d}) \cdot \mathsf{P}(d(R)=t) \\
\iff\ & \sum_{t=0}^{\bar{d}} (\bar{d}-t) \cdot \mathsf{P}(d(R)=t) = \sum_{t=\bar{d}+1}^{\infty} (t-\bar{d}) \cdot \mathsf{P}(d(R)=t) \\
\iff\ & \mathsf{E}[(\bar{d}-d(R))^+] = \mathsf{E}[(d(R)-\bar{d})^+].
\end{aligned} \tag{9}$$

where $(.)^+ := \max\{., 0\}$. It follows from (9) that if C is such that $\mathsf{E}[(C - d(R))^+] \geq \mathsf{E}[(d(R) - C)^+]$ then $\bar{d} \leq C$. Hence, it suffices to show that if $Q_p(d(R)) \geq C(1 - \frac{q}{p})$, for some $p \in (0,1)$, then $\mathsf{E}[(C \quad d(R))^+] \quad \mathsf{E}[(d(R) \quad C)^+]$. Since $\mathsf{E}[d(R)jd(R) > C] \quad C(1+q)$, we have that

$$
\begin{aligned}
\mathsf{E}[(d(R) \quad C)^+] &= \sum_{t=C+1}^{\infty} (t \quad C) \;\; \mathsf{P}(d(R)=t) \\
&= \quad C \;\; \mathsf{P}(d(R) > C) + \sum_{t=C+1}^{\infty} t \;\; \mathsf{P}(d(R)=tjd(R) > C) \;\; \mathsf{P}(d(R) > C) \\
&= (\quad C + \mathsf{E}[d(R)jd(R) > C])\, \mathsf{P}(d(R) > C) \\
& \quad \epsilon q C.
\end{aligned}
$$

On the other hand, if $Q_p(d(R)) \quad C\left(1 \quad \frac{q}{p}\right)$, then

$$
\begin{aligned}
\mathsf{E}[(C \quad d(R))^+] \quad & \sum_{t=0}^{\lfloor C(1-\frac{q}{p})\rfloor} (C \quad t) \;\; \mathsf{P}(d(R)=t) \\
& \frac{\epsilon q C}{p} \sum_{t=0}^{\lfloor C(1-\frac{q}{p})\rfloor} \mathsf{P}(d(R)=t) \\
& \frac{\epsilon q C}{p} \sum_{t=0}^{Q_p(d(R))} \mathsf{P}(d(R)=t) \\
& \epsilon q C,
\end{aligned}
$$

as desired. $\square$

Using Theorem 10, we can prove the following result based on the *load factors* proposed by Florio et al. (2020). A load factor $f \quad 0$ is simply a parameter which controls the considered vehicle capacity. Thus, if we specify that the load factor parameter is f, then a route R is deterministically feasible if $\bar{d}(R) \quad fC$ and it is chance-constraint feasible if $\mathsf{P}(d(R) > fC) \quad \epsilon$. We denote by X^f (resp. X^f_{CC}) the set X (resp. X_{CC}) where the capacity C is replaced by fC.

Corollary 11. *Suppose that every chance-constraint feasible route R satisfies $\mathsf{E}[d(R)jd(R) > C] \quad C(1+q)$, where $q \quad 0$. Then, the following hold.*

(a) $X_{\mathrm{CC}} \quad X^{1+\ q}$;

(b) $X_{\mathrm{CC}}^{1-(\ q)=(1-\ )} \quad X$.

We end this discussion by mentioning that checking if $\mathsf{E}[d(R)jd(R) > C] \quad C(1+q)$ is satisfied for all chance-constraint feasible routes may be too much to check in advance, but if one puts further restrictions on the probability distributions (for example if the probability distribution does not have a too heavy tail), then one may be able to find sufficient conditions for the condition in Theorem 10 and Corollary 11 to hold. In fact, if $\epsilon \leq 0.5$ and for every route R we have that $d(R)$ follows a symmetric distribution, then it is easy to see that $\mathcal{X}_{\text{CC}} \subseteq \mathcal{X}$.

4. Computational study

The code was written in C++ and we used the libraries SCIP (Maher et al. (2017)) and Lemon (Dezső et al. (2011)). All the experiments were executed with a single thread in a Linux machine with Intel(R) Xeon(R) Gold 6142 CPU with 2.60GHz. For separating the chance-constraint rounded capacity inequalities, we used the separation procedure of Dinh et al. (2018). We used the CVRPSEP package of Lysgaard et al. (2004) for separating the traditional rounded capacity inequalities and implementing the branching rules.

In order to compare the two-stage, chance-constraint and combined approaches, we designed a new set of benchmark instances for the VRPSD. Recall that most methods for solving the VRPSD in the literature assume independence on the demands distribution. Since we do not make such an assumption, we also generated instances where the demands are correlated. Moreover, we developed new instances where the vehicle is expected to fail multiple times per route. In all instances the vehicle capacity is set to $C = 100$ and the set of scenarios S have cardinality $t = 200$. The vertices are positioned uniformly randomly in the box $[0, 100]^2$ and we use n to denote the number of vertices. The graph is complete and the cost of an edge is given by the euclidean distance between its endpoints rounded to the nearest integer. Finally, for each instance $\mathcal{I}$, we choose the number of vehicles k as the smallest positive integer in the interval $[3, +\infty)$ such that our algorithms for both CVRP and CC-VRPSD (with $\epsilon = 0.05$) were able to find a feasible solution within the time limit of 30 minutes.

4.1. Instance Sets

- Independent instances

  Every instance in this set have a name in the format "I-n(n)-LB(LB)-UB(UB)-V(σ)-(q)". (Henceforth, we use (.) as a placeholder for the actual values.) The value $q \in [10]$ simply acts as a counter, meaning that for each *instance configuration*, we generate 10 instances. The instances demands are set as following. For each $i \in V_+$, the mean $\bar{d}_i \in \mathbb{Z}$ is taken according to a discrete uniform distribution on the interval $[LB, UB]$, denoted by $U(LB, UB)$. For each scenario $s \in S$ and for each $i \in V_+$, we select $\tilde{d}_i^s \in \mathbb{Z}$ by rounding to the closest integer a number sampled from the normal distribution $N(\bar{d}_i, \sigma)$. Finally, for each vertex $i \in V_+$, we set $d_i^s = \min\{C, \max\{0, \tilde{d}_i^s\}\}$.

- High failure instances

Here we are trying to design instances that would fail a lot, while still satisfying the requirement that a route is deterministically feasible or chance-constraint feasible. Thus, we aim to achieve a high number of back and forth trips to the depot to occur, possibly multiple times per route. In order to do so, for a fraction β of the generated scenarios, we will sum a random value taken from the interval $[40, 60]$.

All instances in this set have names in the format "HF-n(n)-LB(LB)-UB(UB)-V(σ)-B(β)-(q)", where $\beta \in \{0.05, 0.2\}$ and for each value of β, q varies from 1 to 10. Fix an instance configuration (including β) and let the set of scenarios S be $\{s_1, \ldots, s_t\}$. We set the scenarios demands of an instance as follows. First, we compute an initial demand vector $\hat{d}$, where for each $i \in V_+$, $\hat{d}_i$ is taken from $U(LB, UB)$. For each scenario $s_j \in S$ and $i \in V_+$, $\tilde{d}_i^{s_j,1}$ is obtained by rounding a number sampled from $N(\hat{d}_i, \sigma)$, while $\tilde{d}_i^{s_j,2}$ is taken from $U(40, 60)$. Next, if $\beta = 0.2$, we set $\tilde{d}_i^{s_j} = \tilde{d}_i^{s_j,1} + \lambda_i^j \tilde{d}_i^{s_j,2}$, where λ_i^j is i.i.d. and follows a Bernoulli distribution where the probability of success ($\lambda_i^j = 1$) is 0.2. If $\beta = 0.05$, to make increases on demand more significant, for each $s_j \in S$ and $i \in V_+$, we set $\tilde{d}_i^{s_j} = \tilde{d}_i^{s_j,1} + \tilde{d}_i^{s_j,2}$, if $j \in [1, \beta t] \cap \mathbb{Z}$, and $\tilde{d}_i^{s_j} = \tilde{d}_i^{s_j,1}$, otherwise. Therefore, for some chosen scenarios, demand is increased for every customer. Finally, for each vertex $i \in V_+$ and $s_j \in S$, we set $d_i^{s_j} = \min\{C, \max\{0, \tilde{d}_i^{s_j}\}\}$.

- Correlated instances

  The first type of instances in this set have names in the format "Corr-n(n)-LB*(LB)*-UB*(UB)*-(*type*)-(q)", with $q \in [10]$. In these instances we only allow positive correlation between the customer demands. We select the means of the demands $\bar{d}$ according to a discrete uniform distribution $U(LB, UB)$ and generate a covariance matrix Σ according to the method of Dinh et al. (2018). The parameter *type* $\in \{\text{Low}, \text{High}\}$ refers to the the type of correlation we consider (see Dinh et al. (2018)). We then generate each scenario demand $\tilde{d}^s$ by following the multivariate normal distribution with mean $\bar{d}$ and covariance matrix Σ. As usual, for each vertex $i \in V_+$ and $s \in S$, we set $d_i^s = \min\{C, \max\{0, \tilde{d}^s\}\}$. The second type of instances in this set have names in the form "Corr-neg-n(n)-LB*(LB)*-UB*(UB)*-(*type*)-(q)", with $q \in [10]$. In order to generate these we take the same steps as in the previous correlated instances, except that, for each pair of distinct customers i, j, we multiply the entries (i, j) and (j, i) of the covariance matrix by -1 with 10% probability.

4.2. Computational results

In this section, we discuss the computational results obtained in our experiments. We used 5% and 20% for the parameter ϵ and we denote the corresponding chance-constraint and combined algorithms by CC-ϵ and Combined-ϵ. First, we show in Table 1 the number of solved instances in the time limit of 30 minutes for each approach and instance set. Column

“All” shows the number of instances that were solved by all methods. Notice that, compared to the two-stage approach, more instances were solved to optimality using the new combined method. A possible explanation for these results might be related with Corollary 11, since for small values of ϵ, both load factors of $1+\epsilon q$ and $1 \quad \frac{q}{1-}$ might approximate the value of 1. Thus, for small values of ϵ, we might have that $X_{\mathrm{CC}}(\epsilon) \quad X$, and then the integer L-shaped method — used to compute the recourse cost — could end up enumerating a smaller number of feasible solutions when solving the combined model.

	Determ.	Two-Stage	CC-0.05	CC-0.20	Comb.-0.05	Comb.-0.20	All
I	**80 / 80**	55 / 80	65 / 80	73 / 80	64 / 80	63 / 80	54 / 80
HF	**30 / 30**	10 / 30	22 / 30	30 / 30	17 / 30	16 / 30	9 / 30
Corr	40 / 40	40 / 40	40 / 40	40 / 40	40 / 40	40 / 40	40 / 40

Table 1: Number of instances solved by each method in each instance set.

Considering only the instances which were solved by all methods, we computed the averages of the first-stage cost, second-stage cost, total cost and failure ratio. The *failure ratio* of a solution $\bar{x}$ is computed with the formula $\max_{R\in\mathcal{R}(\bar{x})} \mathsf{P}(d(R) > C)$. We grouped the results by instance configurations in Table 2 and by instance sets in Table 3. We highlight in boldface when an approach attains the smallest entry for a measurement.

As one should expect, Tables 2 and 3 show that the two-stage method usually achieves the lowest values with respect to the total cost of a solution. However, solutions produced by the two-stage method might have a higher failure ratio than both combined and chance-constraint methods. On the other hand, among all 3 approaches, the combined method (with $\epsilon = 0.05$) achieves lower values with respect to both second-stage cost and failure ratio.

In order to further analyze the solution quality of the different approaches for the VRPSD, we generated Figures 2-4 for the independent instances. (We also generated these for the other instance sets, but the graphs are similar to the independent instances.) We now focus on Figure 2, as the other figures are generated similarly. All the instances considered in Figure 2 are feasible and were solved by all of the methods (column “All” of Table 1). Let I be such an instance. We use the deterministic method as a reference, so to this end, let $z_{\text{DET}}(I)$ be the total cost (including the recourse cost) of an optimal solution to the deterministic model. Consider one of the discussed approaches for handling stochasticity, say the two-stage model. Let $z_{2\mathrm{S}}(I)$ be the total cost of an optimal solution for the two-stage model. The *relative total cost* of two-stage with respect to instance I is given by the expression $z_{2\mathrm{S}}(I)/z_{\text{DET}}(I)$. Figure 2 has a curve corresponding to the two-stage model and a point (x, y) in this curve indicates that a fraction x of the considered instances have a total cost ratio of at most y. In this sense, Figure 2 is similar to a traditional performance profile (Dolan & Moré (2002)). Figures 3 and 4 are generated in a similar

Table 2: Averages values per instance type.

		Total Cost				2nd Cost				Failure Ratio			
Instance	ϵ	Det.	CC	2S	Comb.	Det.	CC	2S	Comb.	Det.	CC	2S	Comb.
I-NEW-n16-LB0-UB20-V1	0.05	428.24	423.55	**421.89**	423.55	10.42	0.46	1.80	**0.36**	0.17	0.01	0.04	**0.01**
I-NEW-n16-LB0-UB20-V1	0.20	428.24	422.98	**421.89**	422.23	10.42	2.61	1.80	**1.60**	0.17	0.04	0.04	**0.03**
I-NEW-n16-LB0-UB20-V10	0.05	474.93	495.89	**468.01**	495.79	24.36	5.61	14.87	**5.36**	0.26	**0.04**	0.15	**0.04**
I-NEW-n16-LB0-UB20-V10	0.20	488.06	489.13	**480.90**	486.82	28.81	16.75	16.78	**13.82**	0.29	0.13	0.18	**0.12**
I-NEW-n16-LB0-UB30-V1	0.05	502.08	502.69	**496.62**	502.69	16.58	**1.39**	8.02	**1.39**	0.17	**0.01**	0.09	**0.01**
I-NEW-n16-LB0-UB30-V1	0.20	502.08	498.31	**496.62**	498.17	16.58	4.51	8.02	**4.37**	0.17	0.05	0.09	**0.04**
I-NEW-n16-LB0-UB30-V10	0.05	566.34	591.25	**546.85**	591.13	52.12	5.81	23.18	**5.69**	0.37	**0.03**	0.21	**0.03**
I-NEW-n16-LB0-UB30-V10	0.20	566.34	556.04	**546.85**	552.70	52.12	21.60	23.18	**15.14**	0.37	0.15	0.21	**0.13**
I-NEW-n25-LB0-UB20-V1	0.05	532.27	527.86	**524.61**	527.63	15.52	1.73	4.11	**1.39**	0.33	0.03	0.08	**0.02**
I-NEW-n25-LB0-UB20-V1	0.20	532.27	525.21	**524.61**	524.92	15.52	3.71	4.11	**2.80**	0.33	0.08	0.08	**0.06**
I-NEW-n25-LB0-UB20-V10	0.05	512.85	559.79	**500.34**	559.55	33.85	7.79	16.34	**7.55**	0.42	**0.05**	0.17	**0.05**
I-NEW-n25-LB0-UB20-V10	0.20	512.85	**500.34**	**500.34**	**500.34**	33.85	**16.34**	**16.34**	**16.34**	0.42	**0.17**	**0.17**	**0.17**
I-NEW-n25-LB0-UB30-V1	0.05	671.58	659.61	**656.98**	659.01	24.46	1.11	3.98	**0.38**	0.27	0.01	0.04	**0.01**
I-NEW-n25-LB0-UB30-V1	0.20	671.58	663.99	**656.98**	**656.98**	24.46	12.12	**3.98**	**3.98**	0.27	0.13	**0.04**	**0.04**
I-NEW-n25-LB0-UB30-V10	0.05	-	-	-	-	-	-	-	-	-	-	-	-
I-NEW-n25-LB0-UB30-V10	0.20	-	-	-	-	-	-	-	-	-	-	-	-
HF-n15-LB10-UB20-V5-B0.2	0.05	721.14	872.20	**716.48**	872.20	93.64	**21.45**	72.48	**21.45**	0.44	**0.05**	0.34	**0.05**
HF-n15-LB10-UB20-V5-B0.2	0.20	690.83	711.10	**687.10**	711.10	86.83	**47.30**	69.90	**47.30**	0.37	**0.17**	0.29	**0.17**
HF-n15-LB10-UB20-V5-B0.05	0.05	506.85	574.79	**505.01**	574.69	55.25	48.79	53.21	**48.49**	0.11	**0.05**	0.11	**0.05**
HF-n15-LB10-UB20-V5-B0.05	0.20	506.85	507.09	505.01	**504.61**	55.25	56.09	**53.21**	53.41	**0.11**	0.14	**0.11**	0.13
HF-n20-LB10-UB20-V5-B0.05	0.05	-	-	-	-	-	-	-	-	-	-	-	-
HF-n20-LB10-UB20-V5-B0.05	0.20	-	-	-	-	-	-	-	-	-	-	-	-
Corr-n16-LB5-UB33-High	0.05	**609.87**	**609.87**	**609.87**	**609.87**	**2.69**	**2.69**	**2.69**	**2.69**	**0.02**	**0.02**	**0.02**	**0.02**
Corr-n16-LB5-UB33-High	0.20	**609.87**	**609.87**	**609.87**	**609.87**	**2.69**	**2.69**	**2.69**	**2.69**	**0.02**	**0.02**	**0.02**	**0.02**
Corr-n16-LB5-UB33-Low	0.05	605.83	**604.80**	**604.80**	**604.80**	4.63	**1.00**	**1.00**	**1.00**	0.05	**0.01**	**0.01**	**0.01**
Corr-n16-LB5-UB33-Low	0.20	605.83	604.98	**604.80**	**604.80**	4.63	1.98	**1.00**	**1.00**	0.05	0.02	**0.01**	**0.01**
Corr-neg-n16-LB5-UB33-High	0.05	622.22	**622.17**	**622.17**	**622.17**	3.22	**2.67**	**2.67**	**2.67**	0.02	**0.02**	**0.02**	**0.02**
Corr-neg-n16-LB5-UB33-High	0.20	622.22	622.22	**622.17**	**622.17**	3.22	3.22	**2.67**	**2.67**	0.02	0.02	**0.02**	**0.02**
Corr-neg-n16-LB5-UB33-Low	0.05	604.43	604.22	**603.36**	604.22	3.98	**0.00**	1.70	**0.00**	0.07	**0.00**	0.02	**0.00**
Corr-neg-n16-LB5-UB33-Low	0.20	604.43	**603.36**	**603.36**	**603.36**	3.98	**1.70**	**1.70**	**1.70**	0.07	**0.02**	**0.02**	**0.02**

	1st Cost	2nd Cost	Total Cost	Failure Ratio
Independent instances				
Deterministic	**500.43**	23.59	524.01	0.26
Two-Stage	504.89	9.14	**514.03**	0.10
CC-0.05	525.87	2.61	528.48	**0.02**
CC-0.20	508.69	9.56	518.24	0.09
Combined-0.05	525.94	**2.38**	528.32	**0.02**
Combined-0.20	508.69	9.60	518.24	0.09
High failure instances				
Deterministic	**529.78**	72.31	602.09	0.25
Two-Stage	537.22	61.77	**599.06**	0.21
CC-0.05	670.33	36.64	706.97	**0.05**
CC-0.20	564.11	50.04	614.15	0.16
Combined-0.05	670.44	**36.47**	706.92	**0.05**
Combined-0.20	564.22	48.55	612.78	0.15
Correlated instances				
Deterministic	**607.13**	3.60	610.72	0.04
Two-Stage	608.18	2.04	**610.21**	0.02
CC-0.05	608.75	**1.66**	610.41	**0.01**
CC-0.20	607.85	2.42	610.27	0.02
Combined-0.05	608.75	**1.66**	610.41	**0.01**
Combined-0.20	608.18	2.04	**610.21**	0.02

Table 3: Comparisons between different methods for SVRP.

fashion. Namely, for each NAME $\in \{\text{Det}, \text{2S}, \text{CC-0.05}, \text{Comb.-0.05}, \text{CC-0.2}, \text{Comb.-0.2}\}$, we define $r_{\text{NAME}}(I)$ and $f_{\text{NAME}}(I)$ as being the recourse cost and failure ratio of an optimal solution $\bar{x}$ to the model given by NAME. Then, the *relative recourse cost* and *relative failure ratio* of a method NAME is given by $r_{\text{NAME}}(I)/r_{\text{DET}}(I)$ and $f_{\text{NAME}}(I)/f_{\text{DET}}(I)$, respectively. We remark that, in our experiments, whenever the denominators $r_{\text{DET}}(I)$ and $f_{\text{DET}}(I)$ are zero, the numerators $r_{\text{NAME}}(I)$ and $f_{\text{NAME}}(I)$ are also zero (for all choices of NAME), and in such cases, we consider that the relative recourse cost and the relative failure ratio have both a value of 1.

Similarly to Table 3, Figure 2 shows that, among all the three methods, the two-stage model achieves the best relative total costs. However, for approximately 80% of the instances, both combined and chance-constraint methods, with $\epsilon = 0.2$, exhibit similar performance as two-stage. When $\epsilon = 0.05$, Figure 2 shows that, in order to satisfy the stricter requirements, the total cost of a solution increases up to 5% for approximately 83% of the instances, and for all the instances it increases up to 15%. Next, we turn to Figures 3 and 4. Observe that, compared to two-stage, both combined methods and CC-0.05 achieve lower relative recourse cost and relative failure ratio for all the instances. This motivates these two methods for applications where the decision maker is less tolerant to unpredictable costs. In such scenarios, one may prefer solutions with slightly higher total costs, if such solutions are more robust to undesirable uncertainties. Furthermore, our experiments show that, for each $\epsilon \in \{0.05, 0.2\}$, Combined-$\epsilon$ dominates CC-ϵ with respect to all the considered ratios. Thus, if one is sufficiently risk averse to use a chance-constraint model, then it might be better to use the combined approach instead.

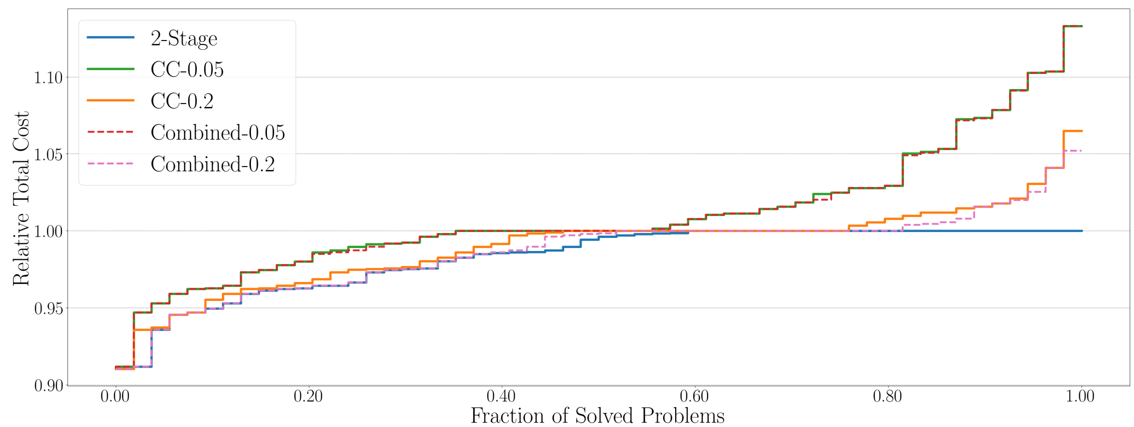

Figure 2: Relative total cost for independent instances.

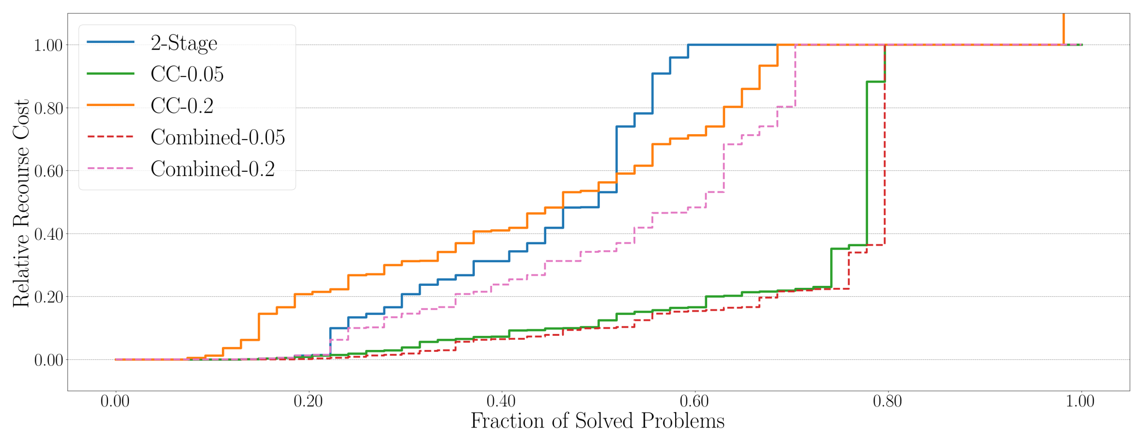

Figure 3: Relative recourse cost for independent instances.

Observe that Figure 2 shows a small difference between the costs of optimal solutions to CC-VRPSD and 2SCC-VRPSD. Our experiments show that, optimal solutions (for any of the approaches) typically execute a recourse action at most once per route and per scenario. Thus, the small difference between CC-VRPSD and 2SCC-VRPSD is in line with the bound shown in Theorem 5. However, while the total overall costs are similar, the combined approach usually obtains solutions with smaller second-stage costs. Therefore, it seems that the combined approach is trading-off first-stage costs for second-stage costs.

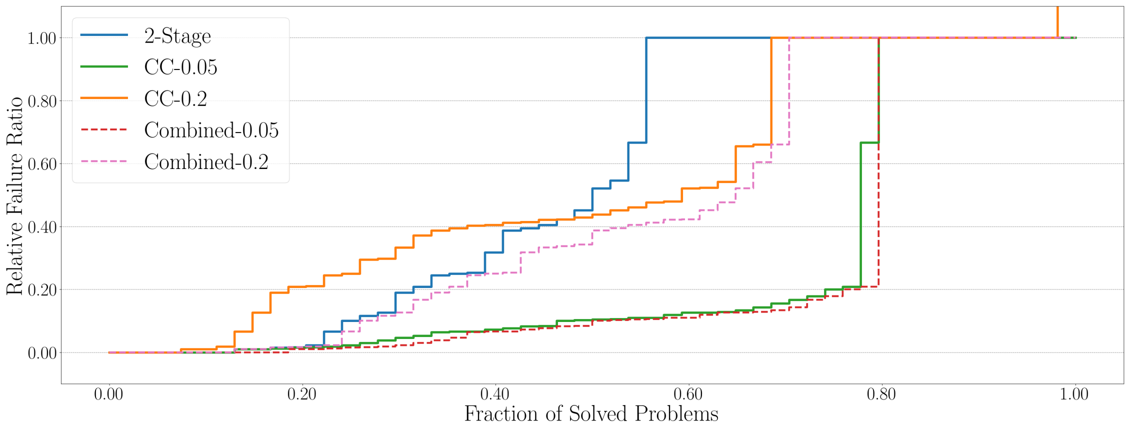

Figure 4: Relative failure ratio for independent instances.

We then decided to take a closer look at this last observation. From our perspective, there can be situations where the weight of a second-stage cost may be higher than a first-stage cost. For instance, such situations can arise when the service provider pays a significant penalty for violating the total capacity of the vehicle, or for deviating from a planned route. Thus, we considered multiplying the second-stage cost by a factor $\gamma \in \mathbb{R}_+$, i.e., we replaced $Q(x)$ in the objective function by $Q\ (x) := \gamma\ \ Q(x)$. We decided to investigate what would happen in these cases and so we executed experiments in the independent instances with $\gamma = 10$. Unfortunately, our algorithms for two-stage and combined VRPSD were not able to solve most of these instances to optimality within the time limit. Regardless, we considered the best primal bound found by the algorithm and generated Figures 5, 6 and 7. We see that in this setting, for $\epsilon = 0.2$, the combined method achieves a significantly smaller total cost than the chance-constraint approach, while still obtaining lower failure ratios than the two-stage method; this is in line with our initial hypothesis that when second-stage costs are higher than first-stage ones, the combined approach seems to be better. Of course, this advantage is expected to diminish once ϵ starts getting smaller, which is confirmed by the experiments shown in Figures 5, 6 and 7 for $\epsilon = 0.05$.

4.2.1. Trade-offs between cost and failure ratio

Intuitively, one would expect that total cost and failure ratio are conflicting objectives. Indeed, the previous experiments show that there are trade-offs between minimizing/constraining each of these measurements. However, the values of ϵ chosen for those experiments were somewhat arbitrary, based solely on two values that we deemed to be reasonable values of ϵ one would be interested in (or at least, to be representative values to show what would happen given a small or high tolerance for risk). Seeking a more complete picture of how these parameters influenced each other, and inspired by the multi-objective optimization literature, we computed a Pareto frontier (Figure 8) for the independent instances using

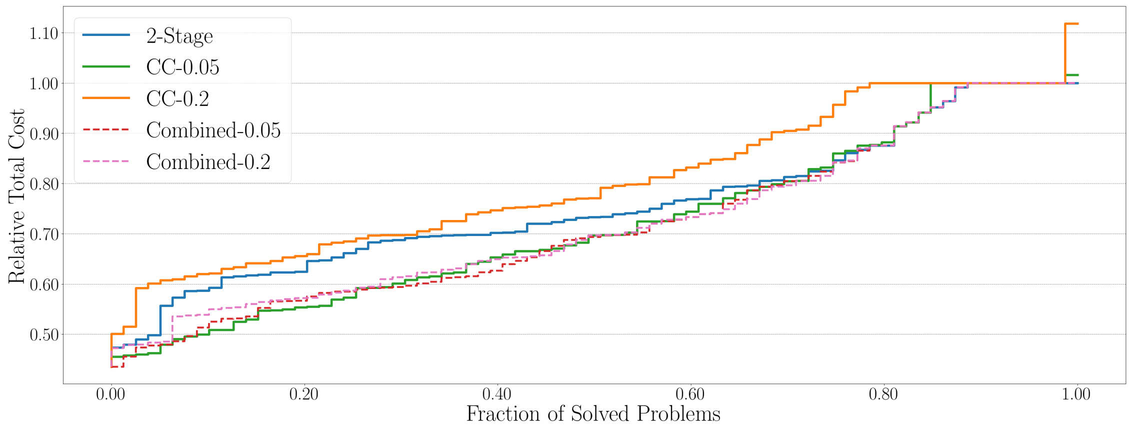

Figure 5: Relative total cost for independent instances with $\gamma = 10$.

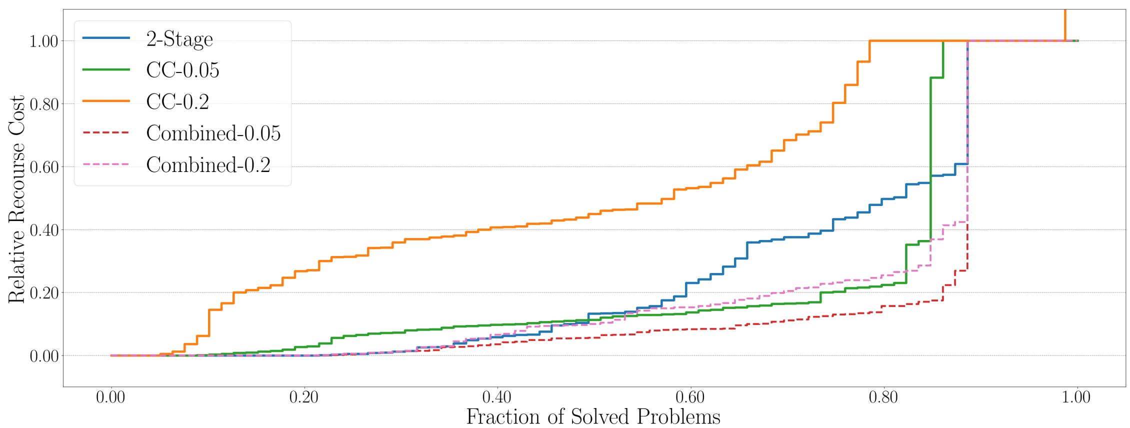

Figure 6: Relative recourse cost for independent instances with $\gamma = 10$.

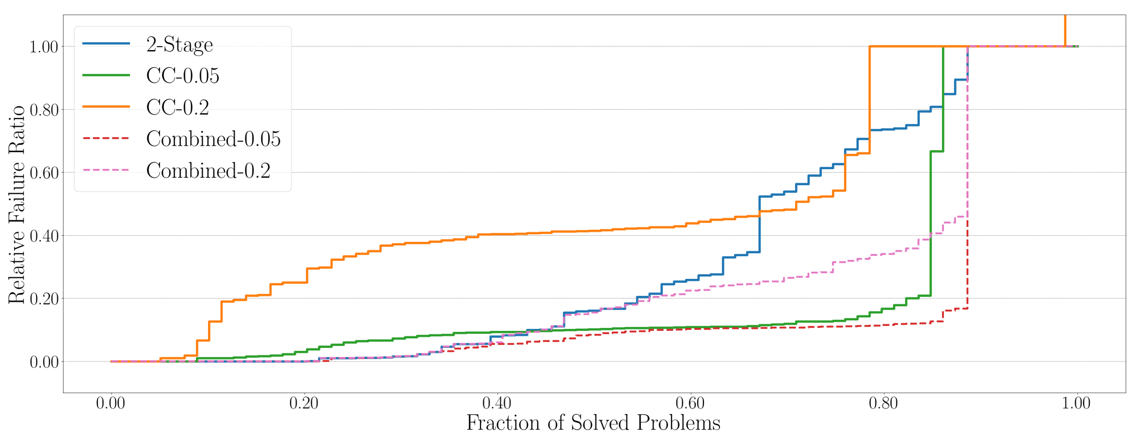

Figure 7: Relative failure ratio for independent instances with $\gamma = 10$.

Algorithm 1.

For each instance, Algorithm 1 keeps track of how changes in the tolerance parameter ϵ

affects the total cost of both chance-constraint and combined methods. For presentation purposes, we denote by $L_{\text{CC}}(\mathcal{I})$ and $L_{\text{2SCC}}(\mathcal{I})$ the output of Algorithm 1 on instance $\mathcal{I}$. Consider an iteration of the while loop (lines 5-11) and let $\epsilon' \in [\epsilon^*, \epsilon]$. Note that $x^*_{\text{CC}}(\mathcal{I}, \epsilon)$ is an optimal solution for the optimization problem solved by CC-SVRP$(\mathcal{I}, \epsilon')$, since $x^*_{\text{CC}}(\mathcal{I}, \epsilon) \in \mathcal{X}_{\text{CC}}(\epsilon^*) \subseteq \mathcal{X}_{\text{CC}}(\epsilon')$. In this sense, we extend the notation so that $z_{\text{CC}}(\mathcal{I}, \epsilon') = z_{\text{CC}}(\mathcal{I}, \epsilon)$ and $z_{\text{2SCC}}(\mathcal{I}, \epsilon') = z_{\text{2SCC}}(\mathcal{I}, \epsilon)$.

Algorithm 1 Algorithm for generating the Pareto frontier.

Input: An instance $\mathcal{I} = (G = (V, E), c, p, k, C)$ of the SVRP.

Output: A list L_{CC} (resp. L_{2SCC}) where an entry is of the form (ϵ, z, x^*), and z denotes the optimal cost to the chance-constraint (resp. combined) model with tolerance parameter ϵ.

Routine: Exact algorithms CC-SVRP() and 2SCC-VRP() which solve the chance-constraint and combined SVRP, respectively.

```
1:  procedure GenerateParetoFrontier(I)
2:      L_CC ← ∅
3:      L_2SCC ← ∅
4:      ε ← 0.5
5:      while ε ≥ 0.05 and X_CC(ε) ≠ ∅ do
6:          Solve CC-SVRP(I, ε) and 2SCC-SVRP(I, ε).
7:          Let x*_CC(I, ε) and x*_2SCC(I, ε) be optimal solutions of CC-SVRP(I, ε)
            and 2SCC-SVRP(I, ε), respectively.
8:          L_CC ← L_CC ∪ {(ε, z_CC(I, ε) = c^T x*_CC(I, ε) + Q(x*_CC(I, ε)))}
9:          L_2SCC ← L_2SCC ∪ {(ε, z_2SCC(I, ε) = c^T x*_2SCC(I, ε) + Q(x*_2SCC(I, ε)))}
10:         Let ε* be the failure ratio of x*_CC(I, ε).
11:         ε ← ε* − 0.001
12:     return L_CC, L_2SCC
```

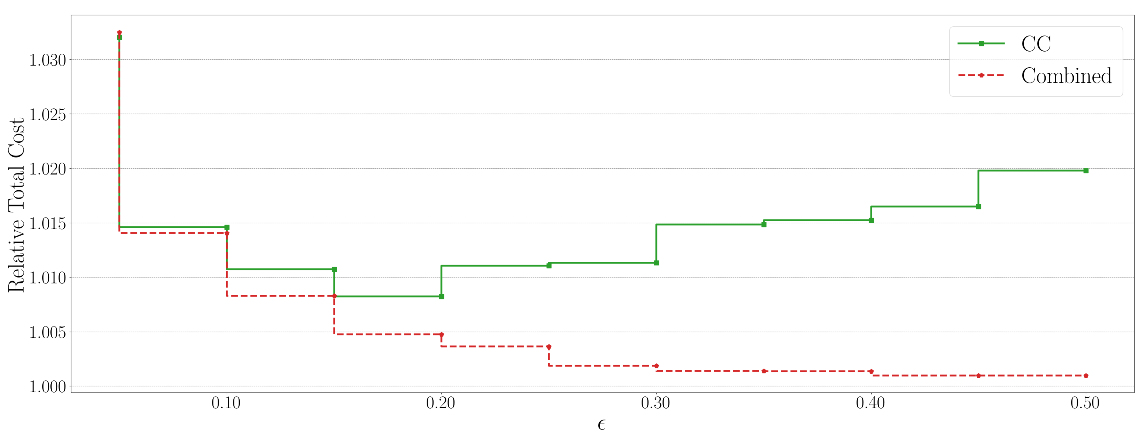

Figure 8: Pareto frontier graph for the independent instances.

Our implementation of Algorithm 1 called the algorithms in line 6 with a time limit of 30 minutes, and for some instances and tolerance parameters, the algorithms were not solved to optimality. In order to cope with this, we implemented a simple preprocessing step that removes some of the instances from consideration. Let $S'_{\mathcal{I}}$ be the set of independent instances $\mathcal{I}$ such that every call to CC-SVRP($\mathcal{I}, \epsilon$) in Algorithm 1 was able to find an optimal solution within the time limit, and whenever the condition in line 5 is evaluated, we have that $X_{\mathrm{CC}}(\epsilon) = \emptyset$. Let $S''_{\mathcal{I}}$ be the set of independent instances for which the two-stage approach was not able to find an optimal solution within the time limit. The set of instances considered in Figure 8 is denoted by $S_{\mathcal{I}}$ and is given by the expression $S'_{\mathcal{I}} \setminus S''_{\mathcal{I}}$. In our experiments, $|S_{\mathcal{I}}| = 52$.

We now detail how Figure 8 was generated. Fix ϵ to be a value in $\{0.05 \, q : q \in [10]\}$. Recall that $z_{2\mathrm{S}}(\mathcal{I})$ is the optimal objective value of the two-stage method in instance $\mathcal{I}$. Similarly to how we computed the relative total cost previously, for a given instance $\mathcal{I} \in S_{\mathcal{I}}$ and method $\textsc{name} \in \{\mathrm{CC}, 2\mathrm{SCC}\}$, we computed the fraction $z_{\textsc{name}}(\mathcal{I}, \epsilon)/z_{2\mathrm{S}}(\mathcal{I})$. In Figure 8, for each $\epsilon \in \{0.05 \, q : q \in [10]\}$ and $\textsc{name} \in \{\mathrm{CC}, 2\mathrm{SCC}\}$, we show the value of the expression

$$\frac{1}{|S_{\mathcal{I}}|} \sum_{\mathcal{I} \in S_{\mathcal{I}}} \frac{z_{\textsc{name}}(\mathcal{I}, \epsilon)}{z_{2\mathrm{S}}(\mathcal{I})}.$$

Observe that, in average, both chance-constrained and combined methods have cost ratios larger than 1, i.e., they output solutions with higher total costs when compared to optimal solutions for the two-stage method. Moreover, enforcing the chance-constraints with $\epsilon = 0.05$ increase the total costs in approximately 3.2% (in average) when in comparison to an optimal solutions for two-stage. If we allow a higher failure ratio and set $\epsilon = 0.2$, the combined method returns solutions which has a modest increase (w.r.t. 2S-VRPSD) of approximately 0.5% (in average). As the tolerance ϵ gets larger, the difference between the costs of combined and chance-constraint methods becomes more significant. This leads us to conclude that, if the practitioner wants to enforce chance-constraints for a probability that is larger than 10%, then using the combined method might lead to considerable savings, when in comparison to the chance-constraint approach. In addition, one might initially expect that the total cost of a solution would decrease as one increases ϵ. However, since CC-VRPSD does not explicitly consider the recourse cost in the objective function, as we increase ϵ, the total cost of an optimal solution might increase or decrease. In this sense, the combined method might be more appropriate, since its curve is always non-increasing in ϵ.

5. Conclusion

In this paper, we compared three different methods for solving the VRPSD. Two of the methods were already studied in the literature, namely, the two-stage and the chance-constraint approach. The third method, a combined approach, was not studied before and was introduced in this work. We conducted comparisons from both a theoretical as from a practical perspective. From a theoretical side, we derived new bounds that show how much costlier can an optimal solution be when it ignores the recourse costs in the objective function. Our results show that an optimal solution for the deterministic model can be considerably more expensive than a solution to the two-stage model. On the other hand, the cost of an optimal solution for the chance-constraint model can be at most $t\epsilon$ times the cost of the combined model, where t is the maximum number of failure of a route in the solution. When $t = 1$, we prove that these bounds can be made arbitrarily tight by constructing instances whose graphs resemble a circle. Furthermore, by examining closely the quantiles of the probability distribution of the accumulated demand of a route, we derived sufficient conditions under which chance-constraint feasible routes are deterministically feasible. From the practical side, we conducted extensive computational experiments which corroborate to the bounds derived in our theoretical study. When ϵ is small, we see that the total costs of the chance-constraint and combined approaches are very similar. As ϵ gets larger, the difference between combined and chance-constraint can be more prominent, particularly if the second stage-costs are significantly high. Furthermore, while the two-stage method consistently attains lower total cost, its solutions may have a larger failure ratio than both chance-constraint and combined approaches. On the other hand, the combined approach attains the smaller failure ratios and second-stage costs.

Overall, while it may be hard to pinpoint a single "best" approach to deal with the VRPSD, given that such a determination will vary according to the decision-maker's risk-aversion and desire to balance expected cost with risk, we believe that this work sheds more light in a quantitative and formal way into allowing decision-makers to determine which approach is more suited for their particular application. Still, there are several other studies that would be worth pursuing and that are left as future work. These include considering more complex recourse strategies and trying to adapt some of our results that do not hold for any recourse policy (for instance the examples of almost-tight behaviour). Another line of research that would be worth investigating is to improve on the algorithms for the two-stage variant of the problem considering correlated demands.

Acknowledgments

The authors acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC) (funding reference number RGPIN-2020- 04030).

References

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

Dantzig, G. B., & Ramser, J. H. (1959). The truck dispatching problem. *Management science*, *6*, 80–91.

Dezső, B., Jüttner, A., & Kovács, P. (2011). Lemon–an open source c++ graph template library. *Electronic Notes in Theoretical Computer Science*, *264*, 23–45.

Dinh, T., Fukasawa, R., & Luedtke, J. (2018). Exact algorithms for the chance-constrained vehicle routing problem. *Mathematical Programming*, *172*, 105–138.

Dolan, E. D., & Moré, J. J. (2002). Benchmarking optimization software with performance profiles. *Mathematical programming*, *91*, 201–213.

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

Gauvin, C., Desaulniers, G., & 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., Laporte, G., & Séguin, R. (1995). An exact algorithm for the vehicle routing problem with stochastic demands and customers. *Transportation Science*, *29*, 143–155.

Golden, B. L., Raghavan, S., Wasil, E. A. et al. (2008). *The vehicle routing problem: latest advances and new challenges* volume 43. Springer.

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

Laporte, G., Louveaux, F., & Mercure, H. (1989). Models and exact solutions for a class of stochastic location-routing problems. *European Journal of Operational Research*, *39*, 71–78.

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

Laporte, G., & Nobert, Y. (1983). A branch and bound algorithm for the capacitated vehicle routing problem. *Operations-Research-Spektrum*, *5*, 77–85.

Louveaux, F. (1998). An introduction to stochastic transportation models. In M. Labbé, G. Laporte, K. Tanczos, & P. Toint (Eds.), *Operations Research and Decision Aid Methodologies in Traffic and Transportation Management* (pp. 244–263). Berlin, Heidelberg: Springer Berlin Heidelberg.

Lysgaard, J., Letchford, A. N., & Eglese, R. W. (2004). A new branch-and-cut algorithm for the capacitated vehicle routing problem. *Mathematical Programming*, *100*, 423–445.

Maher, S. J., Fischer, T., Gally, T., Gamrath, G., Gleixner, A., Gottwald, R. L., Hendel, G., Koch, T., Lübbecke, M. E., Miltenberger, M., Müller, B., Pfetsch, M. E., Puchert, C., Rehfeldt, D., Schenker, S., Schwarz, R., Serrano, F., Shinano, Y., Weninger, D., Witt, J. T., & Witzig, J. (2017). *The SCIP Optimization Suite 4.0*. Technical Report 17-12 ZIB Takustr. 7, 14195 Berlin.

Oyola, J., Arntzen, H., & Woodruff, D. L. (2017). The stochastic vehicle routing problem, a literature review, part ii: solution methods. *EURO Journal on Transportation and Logistics*, *6*, 349–388.

Oyola, J., Arntzen, H., & Woodruff, D. L. (2018). The stochastic vehicle routing problem, a literature review, part i: models. *EURO Journal on Transportation and Logistics*, *7*, 193–221.

Salavati-Khoshghalb, M., Gendreau, M., Jabali, O., & Rei, W. (2019a). A hybrid recourse policy for the vehicle routing problem with stochastic demands. *EURO Journal on Transportation and Logistics*, *8*, 269–298.

Salavati-Khoshghalb, M., Gendreau, M., Jabali, O., & Rei, W. (2019b). A rule-based recourse for the vehicle routing problem with stochastic demands. *Transportation Science*, *53*, 1334–1353.

Stewart, W. R., & Golden, B. L. (1983). Stochastic vehicle routing: A comprehensive approach. *European Journal of Operational Research*, *14*, 371–385.

Toth, P., & Vigo, D. (2014). *Vehicle routing: problems, methods, and applications*. SIAM.

Yee, J. R., & Golden, B. L. (1980). A note on determining operating strategies for probabilistic vehicle routing. *Naval Research Logistics Quarterly*, *27*, 159–163.