Approximation Algorithms for Cumulative Vehicle Routing with Stochastic Demands

Jingyang Zhao
University of Electronic Science and Technology of China, Chengdu, China

Mingyu Xiao[1]
University of Electronic Science and Technology of China, Chengdu, China

Abstract

In the Cumulative Vehicle Routing Problem (Cu-VRP), we need to find a feasible itinerary for a capacitated vehicle located at the depot to satisfy customers' demand, as in the well-known Vehicle Routing Problem (VRP), but the goal is to minimize the cumulative cost of the vehicle, which is based on the vehicle's load throughout the itinerary. If the demand of each customer is unknown until the vehicle visits it, the problem is called Cu-VRP with Stochastic Demands (Cu-VRPSD). In this paper, we propose a randomized 3.456-approximation algorithm for Cu-VRPSD, improving the best-known approximation ratio of 6 (Discret. Appl. Math. 2020). Since VRP with Stochastic Demands (VRPSD) is a special case of Cu-VRPSD, as a corollary, we also obtain a randomized 3.25-approximation algorithm for VRPSD, improving the best-known approximation ratio of 3.5 (Oper. Res. 2012). At last, we give a randomized 3.194-approximation algorithm for Cu-VRP, improving the best-known approximation ratio of 4 (Oper. Res. Lett. 2013).

2012 ACM Subject Classification Theory of computation → Approximation algorithms analysis

Keywords and phrases Cumulative Vehicle Routing, Stochastic Demands, Approximation Algorithms

Digital Object Identifier 10.4230/LIPIcs.ISAAC.2024.59

Funding The work is supported by the National Natural Science Foundation of China, under the grants 62372095, 62172077, and 62350710215.

1 Introduction

In the well-known Vehicle Routing Problem (VRP) [7], we are given an undirected complete graph $G = (V, E)$ with $V = \{v_0, v_1, \dots, v_n\}$, where v_0 denotes the depot, and the other n vertices denote n customers. Moreover, there is a weight function w on the edges representing the length of edges, which satisfies the triangle inequality, and a demand vector $d = (d_1, ..., d_n)$ implying that each customer v_i has a demand of d_i. The objective is to determine an *itinerary* for a vehicle with a capacity of Q, starting from and ending at the depot, that fulfills every customer's demand while minimizing the total weight of the edges in the itinerary.

In the Cumulative Vehicle Routing Problem (Cu-VRP) [18, 19], the goal is also to find an itinerary for the vehicle, but with the objective of minimizing the *cumulative cost* of the itinerary. Here, the cumulative cost for the vehicle traveling from u to v carrying a load of $x_{uv} \leq Q$ units of goods is defined as $a \cdot w(u, v) + b \cdot x_{uv} \cdot w(u, v)$, where $a, b \in \mathbb{R}_{\geq 0}$ are given parameters. Cu-VRP captures the fuel consumption in transportation and logistics, as fuel consumption depends on both the weight of the empty vehicle and the weight of the goods being carried by the vehicle [10]. Since fuel consumption can account for as much as 60% of a vehicle's operational costs [23], Cu-VRP has been studied extensively through both experimental algorithms [27, 25, 9, 13, 22] and approximation algorithms [10, 11, 12]. A recent survey of Cu-VRP can be found in [6].

[1] Corresponding author

35th International Symposium on Algorithms and Computation (ISAAC 2024).
Editors: Julián Mestre and Anthony Wirth; Article No. 59; pp. 59:1–59:18
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

In VRP with Stochastic Demands (VRPSD) [2], the demand of each customer is represented by an independent random variable with a known distribution, and its value is unknown until the vehicle visits the customer. The goal is to design a policy such that the expected weight of the itinerary is minimized. Early surveys on this topic can be found in [14, 3]. Cu-VRP with Stochastic Demands (Cu-VRPSD) was proposed in [11], and similarly, the goal is to design a policy such that the expected cumulative cost of the itinerary is minimized. In VRPSD, we can fully load the vehicle before it leaves the depot. However, in Cu-VRPSD, due to the fact that both higher and lower loads can lead to higher cumulative costs, we need to carefully consider how the vehicle is loaded. This property makes Cu-VRPSD both more challenging and more interesting compared to VRPSD.

In each of the above problems, the *splittable* (resp., *unsplittable*) variant requires that the demand of each customer can be satisfied partially within the vehicle's visits (resp., must be satisfied entirely in one of the vehicle's visits).

In this paper, we consider approximation algorithms for the unsplittable variants. Note that the unsplittable variants are more difficult. For example, unsplittable VRP generalizes the bin packing problem even on a line shape graph [26], and thus cannot be approximated with an approximation ratio of less than 1.5 unless P=NP. For Cu-VRP, an algorithm is called a ρ-approximation algorithm if it can output a solution with a cumulative cost of at most $\rho \cdot \mathrm{OPT}$ in polynomial time, where OPT is the cumulative cost of the optimal solution. For Cu-VRPSD, an algorithm is called a ρ-approximation algorithm if it can employ a policy to get a solution with an expected cumulative cost of at most $\rho \cdot \mathrm{OPT}$ in polynomial time, where OPT is the cumulative cost of the minimum expected cumulative cost solution obtained by the optimal policy.

Let α denote the approximation ratio of the metric Traveling Salesman Problem (TSP). It is well-known that $\alpha \leq 1.5$ [5, 24], which is slightly improved to $\alpha \leq 1.5 - 10^{-36}$ [20, 21]. The ratio α for TSP will be frequently used in VRP related problems.

For VRP, there is an $(\alpha + 1)$-approximation algorithm for the splittable case [17], and an $(\alpha + 2)$-approximation algorithm for the unsplittable case [1]. Blauth *et al.* [4] improved the ratio to $\alpha + 1 - \varepsilon$ for the splittable case, and Friggstad *et al.* [8] improved the ratio to $\alpha + 1 + \ln 2 - \varepsilon'$ for the unsplittable case using the LP rounding method, where ε and ε' are small positive constants related to α. Notably, Friggstad *et al.* [8] also gave a combinatorial $(\alpha + 1.75 - \varepsilon')$-approximation algorithm for the unsplittable case. For VRPSD, there is a randomized $(\alpha + 1 + o(1))$-approximation algorithm for the splittable case, and a randomized $(\alpha + Q)$-approximation algorithm for the unsplittable case. Gupta *et al.* [16] improved the ratios to $\alpha + 1$ and $\alpha + 2$, respectively.

For Cu-VRP, Gaur *et al.* [10] proposed a $(1 + \frac{4\alpha}{\sqrt{4\alpha^2+24\alpha+4}-2\alpha})$-approximation algorithm for the splittable case, and a $(1+\frac{4\alpha}{\sqrt{4\alpha^2+24\alpha+4}-(2\alpha+2)})$-approximation algorithm for the unsplittable case. For Cu-VRPSD, Gaur *et al.* [12] gave a randomized $\max\{1+\frac{3}{2}\alpha, 3\}$-approximation algorithm for the splittable case, and a randomized $\max\{2 + \frac{3}{2}\alpha, 6\}$-approximation algorithm for the unsplittable case.

1.1 Our results

In this paper, we focus on the unsplittable cases of Cu-VRPSD, VRPSD, and Cu-VRP, and design improved approximation algorithms for them.

The main idea of the most recent algorithms [16, 12] is as follows. First, we find an α-approximate TSP tour and then the vehicle satisfies customers in the order they appear on the TSP tour. Once the load is less than the serving customer's demand the vehicle goes back to the depot to reload. Our algorithms will also use an α-approximate TSP tour and visit the customers in the order according to the TSP tour. However, we do not strictly satisfy the customers in the order. To reduce the cumulative cost, our vehicle may skip customers with large demands when visiting customers according to the TSP tour (but record their demands) and satisfy them after completing the TSP tour.

Based on the above idea, we propose two novel algorithms for Cu-VRPSD, denoted as $ALG.1(\lambda, \delta)$ and $ALG.2(\lambda, \delta)$.

- In $ALG.1(\lambda, \delta)$, the vehicle will skip customers in $\{v_i \mid d_i > \lambda \cdot Q\}$ and then satisfy each of them by using a single tour;
- In $ALG.2(\lambda, \delta)$, the vehicle will skip customers in $\{v_i \mid d_i > \delta \cdot Q\}$ and then satisfy them either by using a single tour for each or by calling an algorithm for weighted set cover.

Furthermore, in our algorithms, we set upper and lower bounds of the load of the vehicle when traveling along the TSP tour: the load is at least $\delta \cdot Q$ and less than $\lambda \cdot Q$ for some parameters δ and λ. The lower bound can be regarded as the *backup* goods that the vehicle carries. The idea of carrying some backup goods was inspired by a partition algorithm for the TSP tour used for unsplittable VRP [8]. We will show that this approach can reduce the potential cumulative cost caused by visiting customers with demands at most $\delta \cdot Q$ at the expense of increasing the cumulative cost of the vehicle when traveling the TSP tour. So, we need to balance the setting of δ, e.g., we may set $\delta = 0$ when a/b is small.

We will prove that $ALG.1(\lambda, \delta)$ can be used to obtain a randomized algorithm with an expected approximation ratio of 10/3 for $a/b \leq 0.375$ and 3.456 for $0.375 < a/b \leq 1.444$, and by using both $ALG.1(\lambda, \delta)$ and $ALG.2(\lambda, \delta)$, we can get a randomized 3.456-approximation algorithm for $a/b > 1.444$. Hence, we get a randomized 3.456-approximation algorithm for Cu-VRPSD.

Note that Cu-VRPSD reduces to VRPSD when $b = 0$, and this corresponds to $a/b = \infty$. As a corollary, for VRPSD, we also obtain a randomized 3.25-approximation algorithm using the randomized 3.456-approximation algorithm for Cu-VRPSD with $a/b > 1.444$.

For Cu-VRP, we also give two algorithms, denoted as $ALG.3(\lambda, \delta)$ and $ALG.4(\lambda)$. Since the demands of customers are known in advance, in $ALG.3(\lambda, \delta)$, we first obtain a set of tours by applying the randomized rounding method to the LP of weighted set cover, and then satisfy the remaining customers by calling $ALG.1(\lambda, \delta)$; in $ALG.4(\lambda)$, we directly call $ALG.1(\lambda, 0)$. In the tours obtained by calling $ALG.1(\lambda, \delta)$ and $ALG.1(\lambda, 0)$, the load of the vehicle may be greater than the delivered units of goods. So, we also adapt a pre-optimization step to ensure that the load of the vehicle equals the delivered units of goods.

We will show that $ALG.3(\lambda, \delta)$ can be used to obtain a randomized 3.194-approximation algorithm with a running time of $n^{O(\frac{1}{\min\{a/b,1\}})}$ and thus it only works for $a/b > \gamma_0$, where $\gamma_0 > 0$ is any fixed constant, and $ALG.4(\lambda, \delta)$ can be used to obtain a randomized 3.163-approximation algorithm for $a/b < 0.428$. Hence, we get a randomized 3.194-approximation algorithm for Cu-VRP.

A summary of our results under $\alpha = 1.5$ can be found in Table 1. Although our algorithms are simple and neat, the analysis is technically involved. Some parts also need careful calculation. To avoid distraction from our main discussions and also due to the limited space, the proofs of lemmas and theorems marked with "*" are omitted.

2 Notations

In Cu-VRP, we use $G = (V, E)$ to denote the input complete graph, where $V = \{v_0, \ldots, v_n\}$. There is a non-negative weight function $w : E \to \mathbb{R}_{\geq 0}$ on the edges, where $w(u, v)$ denotes the length of edge $uv \in E$. We assume that w is a *metric*, i.e., it is symmetric and satisfies the

Table 1 A summary of the previous approximation ratios and our approximation ratios.

	Previous Results	Our Results
Cu-VRPSD	6 [12]	**3.456**
VRPSD	3.5 [16]	**3.25**
Cu-VRP	4 [10]	**3.194**

triangle inequality. Let $V' \coloneqq V \setminus \{v_0\}$. There is also a demand vector $d = (d_1, ..., d_n) \in \mathbb{R}^{V'}_{[0,Q]}$, where $Q \in \mathbb{R}_{>0}$ is the capacity of the vehicle, and each customer v_i has a required demand $d_i \in [0, Q]$. We let $l_i \coloneqq w(v_0, v_i)$ and $[i] \coloneqq \{1, 2, ..., n\}$.

In Cu-VRPSD, the demand of each customer v_i is represented by an independent random variable $\chi_i \in [0, Q]$, where the distribution of χ_i is usually assumed to be known in advance [2]. Let $\chi = (\chi_1, ..., \chi_n)$, where we assume that χ_i is not identically zero, as v_i can be ignored in such a case. Consequently, any feasible policy must visit every customer at least once [16].

For any random variable L, we use $L \sim U[l, r)$ to indicate that L is uniformly distributed over the interval $[l, r)$, where $l < r$.

A *tour* $T = v_0 v_1 \dots v_i v_0$ is a directed simple cycle, which always contains the depot v_0. We use $E(T)$ to denote the set of edges on T, and $V'(T)$ to denote the set of customers on T. Assume that the vehicle carries a load of x_{eT} units of goods when traveling along $e \in E(T)$. The cumulative cost of T is

$$Cu(T) \coloneqq a \cdot \sum_{e \in E(T)} w(e) + b \cdot \sum_{e \in E(T)} x_{eT} \cdot w(e),$$

where $w(T) \coloneqq \sum_{e \in E(T)} w(e)$ is called the weight of T, $Cu_1(T) \coloneqq a \cdot \sum_{e \in E(T)} w(e)$ is called the vehicle cost of T, and $Cu_2(T) \coloneqq b \cdot \sum_{e \in E(T)} x_{eT} \cdot w(e)$ is called the cargo cost of T. An itinerary $\mathcal{T}$ is a set of tours. A *TSP tour* is an undirected cycle that includes all customers and the depot exactly once. The weight of the minimum weight TSP tour is denoted by τ.

2.1 Problem Definitions

▶ **Definition 1 (Cu-VRPSD).** *Given a complete graph $G = (V, E)$, a metric weight function w, a vehicle capacity $Q \in \mathbb{R}_{>0}$, a random demand variable vector $\chi = (\chi_1, ..., \chi_n)$, and two parameters $a, b \in \mathbb{R}_{\geq 0}$, we need to design a policy to find a feasible itinerary $\mathcal{T}$ such that*

- *the vehicle carries at most Q units of goods on each tour $T \in \mathcal{T}$,*
- *the vehicle delivers goods to customers only in $V'(T)$ on each tour $T \in \mathcal{T}$,*
- *the sum of the delivered demand over all tours for each $v_i \in V'$ equals the demand of v_i,*

and $\mathbb{E}[Cu(\mathcal{T})]$ is minimized.

Note that the demand of each customer is unknown until the vehicle visits it. In Cu-VRP, we have $\chi = d$, where d is known in advance. We assume that the deliveries are unsplittable: each customer may be included in multiple tours, but its demand must be satisfied entirely within exactly one of those tours. Moreover, by scaling each customer's demand χ_i to χ_i/Q and adjusting the parameter b to $b \cdot Q$, without loss of generality, we assume that $Q = 1$.

2.2 The Lower bounds

To analyze approximation algorithms, we recall the following lower bound for Cu-VRPSD.

▶ **Lemma 2** ([12]). *For unsplittable Cu-VRPSD, it holds that $\mathbb{E}[Cu(\mathcal{T}^*)] \geq a \cdot \max\{\tau, \sum_{i \in [n]} 2 \cdot \mathbb{E}[\chi_i] \cdot l_i\} + b \cdot \sum_{i \in [n]} \mathbb{E}[\chi_i] \cdot l_i$.*

When $a = 1$ and $b = 0$, the lower bound in Lemma 2 becomes $\max\{\tau, \sum_{i\in[n]} 2 \cdot \mathbb{E}[\chi_i] \cdot l_i\}$, and it was used in analyzing approximation algorithms for VRPSD in [16]. To analyze our algorithms, we use a stronger lower bound that was implicitly used in the proof of Lemma 2.

▶ **Lemma 3** ([12])**.** *For unsplittable Cu-VRPSD with any demand realization vector $d \in \mathbb{R}^{V'}_{[0,1]}$, it holds that $\mathbb{E}[Cu(\mathcal{T}^*) \mid \chi = d] \geq LB := a \cdot \max\{\tau,\ \eta\} + b \cdot 0.5 \cdot \eta$, where $\eta := \sum_{i\in[n]} 2 \cdot d_i \cdot l_i$.*

Lemma 3 is stronger than Lemma 2 since it holds that $\mathbb{E}[\max\{X, Y\}] \geq \max\{\mathbb{E}[X], \mathbb{E}[Y]\}$ for any random variables X and Y by the property of the maximum function.

▶ **Lemma 4.** *An algorithm is a ρ-approximation algorithm for Cu-VRPSD if, for any possible demand realization vector $d \in \mathbb{R}^{V'}_{[0,1]}$, the algorithm conditioned on $\chi = d$ outputs a solution $\mathcal{T}$ with a cumulative cost of $\mathbb{E}[Cu(\mathcal{T}) \mid \chi = d] \leq \rho \cdot LB$.*

Proof. Since it holds that $\mathbb{E}[Cu(\mathcal{T}) \mid \chi = d] \leq \rho \cdot LB \leq \rho \cdot \mathbb{E}[Cu(\mathcal{T}^*) \mid \chi = d]$ for any possible demand realization vector d, we can get that $\mathbb{E}[Cu(\mathcal{T}) \mid \chi] \leq \rho \cdot \mathbb{E}[Cu(\mathcal{T}^*) \mid \chi]$. Therefore, we have $\mathbb{E}[Cu(\mathcal{T})] = \mathbb{E}[\mathbb{E}[Cu(\mathcal{T}) \mid \chi]] \leq \rho \cdot \mathbb{E}[\mathbb{E}[Cu(\mathcal{T}^*) \mid \chi]] = \rho \cdot \mathbb{E}[Cu(\mathcal{T}^*)]$. ◀

Hence, we may frequently analyze our algorithms conditioned on $\chi = d$, where $d \in \mathbb{R}^{V'}_{[0,1]}$ is any possible demand realization. For the sake of analysis, we let $\gamma := a/b$, and $\sigma := \gamma/\eta$. Note that $b = 0$ corresponds to the case where $\gamma = \infty$, which turns out to be easier, as will be shown in Theorem 14. We also define

$$\int_l^r x^t dF(x) := \frac{\sum_{v_i \in V': l < d_i \leq r} 2 \cdot d_i^t \cdot l_i}{\sum_{v_i \in V'} 2 \cdot d_i \cdot l_i}, \qquad \text{where} \quad t \in \{0, 1, 2\}. \tag{1}$$

Note that $\int_0^1 x dF(x) = 1$. Moreover, for any $0 \leq l \leq r$, we have

$$l \cdot \int_l^r x^{t-1} dF(x) < \int_l^r x^t dF(x) \leq r \cdot \int_l^r x^{t-1} dF(x). \tag{2}$$

3 Two Algorithms for Cu-VRPSD

3.1 The first algorithm

In this section, we will introduce our first algorithm, denoted as $ALG.1(\lambda, \delta)$, which can be used to get a 10/3-approximation algorithm for Cu-VRPSD with any $\gamma \in (0, 0.375]$ and a 3.456-approximation algorithm for Cu-VRPSD with any $\gamma \in [0.375, 1.444]$. Here, $\lambda \in (0, 1]$ and $\delta \in [0, \lambda/2]$ are parameters that will be defined later.

Firstly, $ALG.1(\lambda, \delta)$ computes an α-approximate TSP tour T^*, which will be oriented in either clockwise or counterclockwise direction. Assume that $T^* = v_0 v_1 \ldots v_n v_0$ by renumbering the customers following the orientation. Then, the vehicle in $ALG.1(\lambda, \delta)$ tries to satisfy the customers in the order of $v_1 \ldots v_n$ as they appear on T^*, where the parameters λ and δ ensures that the load of the vehicle during its travel on each edge of T^* is at least δ and less than λ. Moreover, among its load, the δ units of goods are regarded as *backup* goods, and the other units of goods are regarded as *normal* goods. Specifically, if the vehicle carries L_{i-1} demand of normal goods during its travel from v_{i-1} to v_i, we have $0 \leq L_{i-1} < \lambda - \delta$ for each $i \in [n+1]$. We say that the vehicle carries $S_{i-1} = (L_{i-1}, \delta)$ units of goods to indicate that it carries L_{i-1} demand of normal goods and δ demand of backup goods. We require that $0 < \lambda \leq 1$ and $0 \leq \delta \leq \lambda - \delta$, i.e., $0 \leq \delta \leq \lambda/2$. When serving a customer, the main strategy is to prioritize using the normal goods first and then consider using the backup goods if the normal goods are insufficient. Conditioned on $\chi = d$, the details are as follows.

Initially, we load the vehicle with $S_0 = (L_0, \delta)$ units of goods at the depot, where $L_0 \sim U[0, \lambda - \delta)$. When the vehicle is about to serve v_i, we assume that it carries $S_{i-1} = (L_{i-1}, \delta)$ units of goods, where $0 \leq L_{i-1} < \lambda - \delta$. Then, we have the following three cases.

Case 1: $d_i \leq L_{i-1}$**.** In this case, the vehicle directly delivers $(d_i, 0)$ units of goods for v_i, and then goes to the next customer. Hence, we have $S_i = (L_i, \delta)$, where $L_i \coloneqq L_{i-1} + \lceil \frac{d_i - L_{i-1}}{\lambda - \delta} \rceil \cdot (\lambda - \delta) - d_i = L_{i-1} - d_i$ since $d_i \leq L_{i-1} < \lambda - \delta$.

Case 2: $L_{i-1} < d_i \leq L_{i-1} + \delta$**.** The vehicle delivers $(L_{i-1}, d_i - L_{i-1})$ units of goods for v_i, goes to the depot to reload $(L_{i-1} + \lceil \frac{d_i - L_{i-1}}{\lambda - \delta} \rceil \cdot (\lambda - \delta) - d_i, d_i - L_{i-1})$ units of goods, and then goes to the next customer. Hence, we have $S_i = (L_i, \delta)$, where $L_i \coloneqq L_{i-1} + \lceil \frac{d_i - L_{i-1}}{\lambda - \delta} \rceil \cdot (\lambda - \delta) - d_i = L_{i-1} + (\lambda - \delta) - d_i$ since $0 < d_i - L_{i-1} \leq \delta \leq \lambda - \delta$.

Case 3: $L_{i-1} + \delta < d_i \leq 1$**.** In this case, we must have $d_i > \delta$.

- **Case 3.1:** $\delta < d_i \leq \lambda$**.** The vehicle goes to the depot to reload $(d_i - L_{i-1} - \delta, 0)$ units of goods, goes to satisfy v_i, then goes to the depot to reload $(L_{i-1} + \lceil \frac{d_i - L_{i-1}}{\lambda - \delta} \rceil \cdot (\lambda - \delta) - d_i, \delta)$ units of goods, goes to customer v_i again (for the sake of analysis), and then goes to the next customer. Hence, we have $S_i = (L_i, \delta)$, where $L_i \coloneqq L_{i-1} + \lceil \frac{d_i - L_{i-1}}{\lambda - \delta} \rceil \cdot (\lambda - \delta) - d_i$.
- **Case 3.2:** $\lambda < d_i \leq 1$**.** Since $L_{i-1} < \lambda - \delta$, we must have $L_{i-1} + \delta < d_i$. Instead of satisfying v_i by returning to the depot to reload as in Case 3.1, the vehicle records its demand, skips it, and goes to the next customer. Hence, we have $S_i = (L_i, \delta)$, where $L_i \coloneqq L_{i-1}$.

After trying to satisfy all customers using the above strategy, due to Case 3.2, there may be still a set of *unsatisfied* customers $\{v_i \mid d_i > \lambda\}$. Then, for each unsatisfied customer v_i, since its demand has been recorded, the vehicle will load exactly d_i units of goods at the depot, go to satisfy v_i, and then return to the depot.

The details of $ALG.1(\lambda, \delta)$ is shown in Algorithm 1.

Compared to the previous strategy in [12], there are two main differences. The first is that we specially handle each customer v_i with $d_i > \lambda$ in Case 3.2. Note that if we satisfy v_i as the method in Case 3.1, since $L_{i-1} + \delta < \lambda$ (we will prove it in Lemma 6), the vehicle must incur two visits to the depot, which will cost too much. The second is that we ensure that the vehicle always carries δ demand of backup goods when traveling along the TSP tour. The advantage is that each customer v_i with $d_i \leq \delta$ incurs at most one visit to the depot while if $\delta = 0$ every customer v_i with $d_i \leq \lambda$ has the potential to incur two visits to the depot. However, since $\delta > 0$ clearly increases the cumulative cost of the vehicle when it travels along the TSP tour, we need to carefully set the value of δ.

Although we require that $\lambda > 0$ in $ALG.1(\lambda, \delta)$, it can be extended to the case of $\lambda = 0$. In this scenario, the vehicle simply travels along the TSP tour with an empty carry to record each customer's demand, and then satisfies each customer within a single tour, as described in Case 3.2. Interestingly, if $a = 0$, this algorithm becomes an exact algorithm for unsplittable Cu-VRPSD, as the cumulative cost is $b \cdot \sum_{i \in [n]} d_i \cdot l_i$, which matches the lower bound LB in Lemma 4. The running time can reach $O(n)$ since all TSP tours have the same performance. However, it may be useless for $a > 0$. Hence, we consider $\lambda > 0$ in the following.

▶ **Lemma 5.** *Unsplittable Cu-VRPSD with $a = 0$ can be solved in $O(n)$ time.*

3.1.1 The analysis

Note that $ALG.1(\lambda, \delta)$ carries $L_0 \sim U[0, \lambda - \delta)$ demand of normal goods initially. Next, we analyze the expected cumulative cost of $\mathcal{T}$ conditioned on $\chi = d$, i.e., $\mathbb{E}[Cu(\mathcal{T}) \mid \chi = d]$.

Algorithm 1 An algorithm for unsplittable Cu-VRPSD ($ALG.1(\lambda, \delta)$).

Input: An instance of unsplittable Cu-VRPSD, and two parameters $\lambda \in (0, 1]$ and $\delta \in [0, \lambda/2]$.
Output: A feasible solution $\mathcal{T}$ to unsplittable Cu-VRPSD.

1: Obtain an α-approximate TSP tour T^* using an α-approximation algorithm for metric TSP, orient T^* in either clockwise or counterclockwise direction, and denote $T^* = v_0v_1v_2 \dots v_nv_0$ by renumbering the customers following the direction.
2: Load the vehicle with $S_0 \coloneqq (L_0, \delta)$ units of goods, including L_0 demand of normal goods and δ demand of backup goods, where $L_0 \sim U[0, \lambda - \delta)$.
3: Initialize $i \coloneqq 1$ and $V^* \coloneqq \emptyset$.
4: **while** $i \leq n$ **do**
5: Go to customer v_i;
6: **if** $d_i \leq L_{i-1}$ **then**
7: Deliver $(d_i, 0)$ units of goods to v_i, and update $S_i \coloneqq (L_i, \delta)$, where $L_i \coloneqq L_{i-1} + \lceil \frac{d_i - L_{i-1}}{\lambda - \delta} \rceil \cdot (\lambda - \delta) - d_i = L_{i-1} - d_i$;
8: **else if** $L_{i-1} < d_i \leq L_{i-1} + \delta$ **then**
9: Deliver $(L_{i-1}, d_i - L_{i-1})$ units of goods to v_i, goes to the depot, load the vehicle with $(L_{i-1} + \lceil \frac{d_i - L_{i-1}}{\lambda - \delta} \rceil \cdot (\lambda - \delta) - d_i, d_i - L_{i-1})$ units of goods, and update $S_i \coloneqq (L_i, \delta)$, where $L_i \coloneqq L_{i-1} + \lceil \frac{d_i - L_{i-1}}{\lambda - \delta} \rceil \cdot (\lambda - \delta) - d_i = L_{i-1} + (\lambda - \delta) - d_i$;
10: **else if** $L_i + \delta < d_i \leq 1$ **then**
11: **if** $\delta < d_i \leq \lambda$ **then**
12: Return to the depot, load the vehicle with $(d_i - L_{i-1} - \delta, 0)$ units of goods, go to customer v_i, and deliver $(d_i - \delta, \delta)$ units of goods to v_i.
13: Return to the depot, load the vehicle with $(L_{i-1} + \lceil \frac{d_i - L_{i-1}}{\lambda - \delta} \rceil \cdot (\lambda - \delta) - d_i, \delta)$ units of goods, go to customer v_i, and update $S_i \coloneqq (L_i, \delta)$, where $L_i \coloneqq L_{i-1} + \lceil \frac{d_i - L_{i-1}}{\lambda - \delta} \rceil \cdot (\lambda - \delta) - d_i$; ▷ The vehicle returns to v_i for the sake of analysis; however, it could directly proceed to v_{i+1}.
14: **else if** $\lambda < d_i \leq 1$ **then**
15: Record v_i's demand, and update $V^* \coloneqq V^* \cup \{v_i\}$ and $S_i \coloneqq (L_i, \delta)$, where $L_i \coloneqq L_{i-1}$;
16: **end if**
17: **end if**
18: $i \coloneqq i + 1$.
19: **end while**
20: Go to the depot.
21: **for** $v_i \in V^*$ **do**
22: Load the vehicle with d_i units of goods, go to customer v_i, and deliver d_i units of goods to v_i;
23: Go to the depot.
24: **end for**

In $ALG.1(\lambda, \delta)$, the vehicle carries $S_{i-1} = (L_{i-1}, \delta)$ units of goods when traveling along the edge $v_{i-1}v_{i \bmod (n+1)}$ of the TSP tour T^*. For each $i \in [n]$, we let $h_i \coloneqq 1$ if $d_i \leq \lambda$, and $h_i \coloneqq 0$ otherwise. We have the following lemma.

▶ **Lemma 6.** *For any $i \in [n+1]$, it holds $L_{i-1} = L_0 + \lceil \frac{\sum_{j=1}^{i-1} h_j \cdot d_j - L_0}{\lambda - \delta} \rceil \cdot (\lambda - \delta) - \sum_{j=1}^{i-1} h_j \cdot d_j$, and moreover, $L_{i-1} \sim U[0, \lambda - \delta)$, conditioned on $\chi = d$.*

Proof. Since $L_0 \sim U[0, \lambda - \delta)$, the lemma holds for $i = 1$. Assume that the equality holds for $i = i' \geq 1$, i.e., $L_{i'-1} = L_0 + \lceil \frac{\sum_{j=1}^{i'-1} h_j \cdot d_j - L_0}{\lambda - \delta} \rceil \cdot (\lambda - \delta) - \sum_{j=1}^{i'-1} h_j \cdot d_j$. Note that we have $0 \leq L_{i'-1} < \lambda - \delta$. Next, we consider $L_{i'}$.

Case 1: $d_{i'} \leq \lambda$. We have $h_{i'} = 1$. By Lines 7, 9, and 13, we have $L_{i'} = L_{i'-1} + \lceil \frac{d_{i'} - L_{i'-1}}{\lambda - \delta} \rceil \cdot (\lambda - \delta) - d_{i'}$. Hence, we have $L_{i'} \geq 0$ and $L_{i'} < \lambda - \delta$. Therefore, we have $L_0 + \lceil \frac{\sum_{j=1}^{i'-1} h_j \cdot d_j - L_0}{\lambda - \delta} \rceil \cdot (\lambda - \delta) - \sum_{j=1}^{i'} h_j \cdot d_j + \lceil \frac{d_{i'} - L_{i'-1}}{\lambda - \delta} \rceil \cdot (\lambda - \delta) \geq 0$ and

$L_0+\lceil\frac{\sum_{j=1}^{i'-1}h_j\cdot d_j-L_0}{\lambda-\delta}\rceil\cdot(\lambda-\delta)-\sum_{j=1}^{i'}h_j\cdot d_j+\lceil\frac{d_{i'}-L_{i'-1}}{\lambda-\delta}\rceil\cdot(\lambda-\delta)<\lambda-\delta$. Alternatively, we have $\lceil\frac{\sum_{j=1}^{i'-1}h_j\cdot d_j-L_0}{\lambda-\delta}\rceil+\lceil\frac{d_{i'}-L_{i'-1}}{\lambda-\delta}\rceil-1<\frac{\sum_{j=1}^{i'}h_j\cdot d_j-L_0}{\lambda-\delta}\leq\lceil\frac{\sum_{j=1}^{i'-1}h_j\cdot d_j-L_0}{\lambda-\delta}\rceil+\lceil\frac{d_{i'}-L_{i'-1}}{\lambda-\delta}\rceil$, and hence $\lceil\frac{\sum_{j=1}^{i'}h_j\cdot d_j-L_0}{\lambda-\delta}\rceil=\lceil\frac{\sum_{j=1}^{i'-1}h_j\cdot d_j-L_0}{\lambda-\delta}\rceil+\lceil\frac{d_{i'}-L_{i'-1}}{\lambda-\delta}\rceil$. Therefore, we have $L_{i'}=L_{i'-1}+\lceil\frac{d_{i'}-L_{i'-1}}{\lambda-\delta}\rceil\cdot(\lambda-\delta)-d_{i'}=L_0+\lceil\frac{\sum_{j=1}^{i'-1}h_j\cdot d_j-L_0}{\lambda-\delta}\rceil\cdot(\lambda-\delta)-\sum_{j=1}^{i'}h_j\cdot d_j+\lceil\frac{d_{i'}-L_{i'-1}}{\lambda-\delta}\rceil\cdot(\lambda-\delta)=L_0+\lceil\frac{\sum_{j=1}^{i'}h_j\cdot d_j-L_0}{\lambda-\delta}\rceil\cdot(\lambda-\delta)-\sum_{j=1}^{i'}h_j\cdot d_j$.

Case 2: $d_i>\lambda$. We have $h_i=0$, and hence $\sum_{j=1}^{i'-1}h_j\cdot d_j=\sum_{j=1}^{i'}h_j\cdot d_j$. By Line 15, we have $L_{i'}=L_{i'-1}=L_0+\lceil\frac{\sum_{j=1}^{i'-1}h_j\cdot d_j-L_0}{\lambda-\delta}\rceil\cdot(\lambda-\delta)-\sum_{j=1}^{i'-1}h_j\cdot d_j=L_0+\lceil\frac{\sum_{j=1}^{i'}h_j\cdot d_j-L_0}{\lambda-\delta}\rceil\cdot(\lambda-\delta)-\sum_{j=1}^{i'}h_j\cdot d_j$.

In both cases, we have $L_{i'}=L_0+\lceil\frac{\sum_{j=1}^{i'}h_j\cdot d_j-L_0}{\lambda-\delta}\rceil\cdot(\lambda-\delta)-\sum_{j=1}^{i'}h_j\cdot d_j$. By induction, the equality holds for any $i\in[n+1]$.

For any $i\in[n+1]$, we have $L_{i-1}=L_0+\lceil\frac{\sum_{j=1}^{i-1}h_j\cdot d_j-L_0}{\lambda-\delta}\rceil\cdot(\lambda-\delta)-\sum_{j=1}^{i-1}h_j\cdot d_j$. Assume that $(\sum_{j=1}^{i-1}h_j\cdot d_j) \bmod (\lambda-\delta)=L'$, which is fixed conditioned on $\chi=d$. We have $L_{i-1}=\lambda-\delta+L_0-L'\in[\lambda-\delta-L',\lambda-\delta)$ when $L_0\in[0,L')$, and $L_{i-1}=L_0-L'\in[0,\lambda-\delta-L')$ when $L_0\in[L',\lambda-\delta)$. The relationship between L_0 and L_{i-1} is bijective. Since $L_0\sim U[0,\lambda-\delta)$, we can also get $L_{i-1}\sim U[0,\lambda-\delta)$, conditioned on $\chi=d$. ◀

Lemma 6 also implies that $0\leq L_{i-1}<\lambda-\delta$ for any $i\in[n+1]$.

▶ **Lemma 7.** *In ALG.1(λ,δ), the expected cumulative cost conditioned on $\chi=d$ during the vehicle's travel from v_{i-1} to v_i is $a\cdot w(v_{i-1},v_i)+b\cdot\frac{\lambda+\delta}{2}\cdot w(v_{i-1},v_i)$.*

Proof. By Lemma 6, the vehicle carries (L_{i-1},δ) units of goods during the vehicle's travel from v_{i-1} to v_i, where $L_{i-1}\sim U[0,\lambda-\delta)$. So, $\mathbb{E}[L_{i-1}+\delta\mid\chi=d]=\int_0^{\lambda-\delta}\frac{x+\delta}{\lambda-\delta}dx=\frac{\lambda+\delta}{2}$. Hence, the expected cumulative cost of the vehicle's travel from v_{i-1} to v_i conditioned on $\chi=d$ is $a\cdot w(v_{i-1},v_i)+b\cdot\mathbb{E}[L_{i-1}+\delta\mid\chi=d]\cdot w(v_{i-1},v_i)=a\cdot w(v_{i-1},v_i)+b\cdot\frac{\lambda+\delta}{2}\cdot w(v_{i-1},v_i)$. ◀

By Line 9 in *ALG*.1(λ,δ), if the vehicle visits v_i carrying $S_{i-1}=(L_{i-1},\delta)$ units of goods, where $L_{i-1}<d_i\leq L_{i-1}+\delta$, it will first satisfy v_i, then proceed to the depot to reload some units of goods, and finally return to the place of v_i. We refer to this process as *one additional visit to v_0*.

By Lines 12 and 13 in *ALG*.1(λ,δ), if the vehicle visits v_i with $d_i\leq\lambda$ carrying $S_{i-1}=(L_{i-1},\delta)$ units of goods, where $L_{i-1}+\delta<d_i$, it will first go to the depot to reload some units of goods, then go to the place of v_i to satisfy v_i, proceed to the depot to reload some units of goods, and finally return to the place of v_i again. We refer to this process as *two additional visits to v_0*.

When the vehicle is about to serve v_i with $d_i\leq\lambda$, it may incur one additional visit or two additional visits to v_0, resulting in some cumulative cost. For each customer v_i with $d_i>\lambda$, By Line 22, the vehicle satisfies v_i using a single tour, which will also be regarded as one additional visit to v_0 for the sake of presentation. Next, we analyze the expected cumulative cost conditioned on $\chi=d$ due to the possible additional visit(s) for each customer v_i.

▶ **Lemma 8.** *Conditioned on $\chi=d$, when serving each customer v_i in ALG.1(λ,δ), the expected cumulative cost of the vehicle due to the possible additional visit(s) to v_0 is*

- $a\cdot\frac{2d_i}{\lambda-\delta}\cdot l_i+b\cdot\frac{(\lambda+\delta)\cdot d_i-d_i^2}{\lambda-\delta}\cdot l_i$ *if $d_i\leq\delta$;*

- $a \cdot \frac{4d_i - 2\delta}{\lambda - \delta} \cdot l_i + b \cdot \frac{d_i^2 + (\lambda-\delta)\cdot d_i}{\lambda - \delta} \cdot l_i$ *if* $\delta < d_i \leq \lambda - \delta$;
- $a \cdot \frac{2d_i + 2\lambda - 4\delta}{\lambda - \delta} \cdot l_i + b \cdot \frac{2d_i^2 - (\lambda+\delta)\cdot d_i + \lambda^2 - \delta^2}{\lambda - \delta} \cdot l_i$ *if* $\lambda - \delta < d_i \leq \lambda$;
- $a \cdot 2 \cdot l_i + b \cdot d_i \cdot l_i$ *if* $\lambda < d_i \leq 1$.

Proof. If $d_i > \lambda$, By Lines 15 and 22, the vehicle incurs one additional visit to v_0, where the vehicle carries d_i units of goods from v_0 to v_i and 0 units of goods from v_i to v_0. So, the expected cumulative cost is $a \cdot 2 \cdot l_i + b \cdot d_i \cdot l_i$. Next, we consider $d_i \leq \lambda$.

By Lemma 6, the vehicle carries (L_{i-1}, δ) units of goods when traveling along $v_{i-1}v_i$ in $ALG.1(\lambda, \delta)$, and it holds that $L_{i-1} \sim U[0, \lambda - \delta)$. Hence, the vehicle incurs one additional visits to v_0 with a probability of $\Pr[d_i - \delta \leq L_{i-1} = x < d_i]$, and incurs two additional visits to v_0 with a probability of $\Pr[0 \leq L_{i-1} = x < d_i - \delta]$. We consider the following three cases.

Case 1: $d_i \leq \delta$. The vehicle incurs at most one additional visit to v_0. If the vehicle incurs one additional visit to v_0, By Line 9, the vehicle carries $(0, \delta - (d_i - L_{i-1}))$ units of goods from v_i to v_0, and $(L_{i-1} + \lceil \frac{d_i - L_{i-1}}{\lambda - \delta} \rceil \cdot (\lambda - \delta) - d_i, \delta) = (L_{i-1} + (\lambda - \delta) - d_i, \delta)$ units of goods from v_0 to v_i. So, the cumulative cost is $a \cdot 2 \cdot l_i + b \cdot (\delta - (d_i - L_{i-1}) + L_{i-1} + (\lambda - \delta) - d_i + \delta) \cdot l_i = a \cdot 2 \cdot l_i + b \cdot (2L_{i-1} - 2d_i + \lambda + \delta) \cdot l_i$. Since $L_{i-1} \sim U[0, \lambda - \delta)$ and $d_i \leq \delta \leq \lambda - \delta$, the expected cumulative cost is

$$\int_0^{\min\{d_i, \lambda-\delta\}} \frac{a \cdot 2 \cdot l_i + b \cdot (2x - 2d_i + \lambda + \delta) \cdot l_i}{\lambda - \delta} dx$$
$$= \int_0^{d_i} \frac{a \cdot 2 \cdot l_i + b \cdot (2x - 2d_i + \lambda + \delta) \cdot l_i}{\lambda - \delta} dx = a \cdot \frac{2d_i}{\lambda - \delta} \cdot l_i + b \cdot \frac{(\lambda + \delta) \cdot d_i - d_i^2}{\lambda - \delta} \cdot l_i.$$

Case 2: $\delta < d_i \leq \lambda - \delta$. The vehicle incurs at most two additional visits to v_0. Similarly, if the vehicle incurs one additional visit to v_0, By Line 9, the vehicle carries $(0, \delta - (d_i - L_{i-1}))$ units of goods from v_i to v_0, and $(L_{i-1} + \lceil \frac{d_i - L_{i-1}}{\lambda - \delta} \rceil \cdot (\lambda - \delta) - d_i, \delta) = (L_{i-1} + (\lambda - \delta) - d_i, \delta)$ units of goods from v_0 to v_i, and the cumulative cost is $a \cdot 2 \cdot l_i + b \cdot (\delta - (d_i - L_{i-1}) + L_{i-1} + (\lambda - \delta) - d_i + \delta) \cdot l_i = a \cdot 2 \cdot l_i + b \cdot (2L_{i-1} - 2d_i + \lambda + \delta) \cdot l_i$. If the vehicle incurs two additional visits to v_0, By Lines 12 and 13, the vehicle carries (L_{i-1}, δ) units of goods from v_i to v_0, $(d_i - \delta, \delta)$ units of goods from v_0 to v_i, $(0, 0)$ units of goods from v_i to v_0, and $(L_{i-1} + \lceil \frac{d_i - L_{i-1}}{\lambda - \delta} \rceil \cdot (\lambda - \delta) - d_i, \delta) = (L_{i-1} + (\lambda - \delta) - d_i, \delta)$ units of goods from v_0 to v_i, and the cumulative cost is $a \cdot 4 \cdot l_i + b \cdot (L_{i-1} + \delta + d_i + 0 + L_{i-1} + (\lambda - \delta) - d_i + \delta) \cdot l_i = a \cdot 4 \cdot l_i + b \cdot (2L_{i-1} + \lambda + \delta) \cdot l_i$. Hence, the expected cumulative cost is

$$\int_{d_i - \delta}^{\min\{d_i, \lambda-\delta\}} \frac{a \cdot 2 \cdot l_i + b \cdot (2x - 2d_i + \lambda + \delta) \cdot l_i}{\lambda - \delta} dx + \int_0^{d_i - \delta} \frac{a \cdot 4 \cdot l_i + b \cdot (2x + \lambda + \delta) \cdot l_i}{\lambda - \delta} dx$$
$$= \int_{d_i - \delta}^{d_i} \frac{a \cdot 2 \cdot l_i + b \cdot (2x - 2d_i + \lambda + \delta) \cdot l_i}{\lambda - \delta} dx + \int_0^{d_i - \delta} \frac{a \cdot 4 \cdot l_i + b \cdot (2x + \lambda + \delta) \cdot l_i}{\lambda - \delta} dx$$
$$= \frac{a \cdot 2 \cdot \delta \cdot l_i + b \cdot \lambda \cdot \delta \cdot l_i}{\lambda - \delta} + \frac{a \cdot 4 \cdot (d_i - \delta) \cdot l_i + b \cdot (d_i^2 + (\lambda - \delta) \cdot d_i - \delta \cdot \lambda) \cdot l_i}{\lambda - \delta}$$
$$= a \cdot \frac{4d_i - 2\delta}{\lambda - \delta} \cdot l_i + b \cdot \frac{d_i^2 + (\lambda - \delta) \cdot d_i}{\lambda - \delta} \cdot l_i.$$

Case 3: $\lambda - \delta < d_i \leq \lambda$. The vehicle incurs at most two additional visits to v_0. If the vehicle incurs one additional visit to v_0, By Line 9, the vehicle carries $(0, \delta - (d_i - L_{i-1}))$ units of goods from v_i to v_0, and $(L_{i-1} + \lceil \frac{d_i - L_{i-1}}{\lambda - \delta} \rceil \cdot (\lambda - \delta) - d_i, \delta)$ units of goods from v_0 to v_i, and the cumulative cost is $a \cdot 2 \cdot l_i + b \cdot (\delta - (d_i - L_{i-1}) + L_{i-1} + \lceil \frac{d_i - L_{i-1}}{\lambda - \delta} \rceil \cdot (\lambda - \delta) - d_i + \delta) \cdot l_i = a \cdot 2 \cdot l_i + b \cdot (2L_{i-1} - 2d_i + 2\delta + \lceil \frac{d_i - L_{i-1}}{\lambda - \delta} \rceil \cdot (\lambda - \delta)) \cdot l_i$. If the vehicle incurs two additional visits to v_0, By Lines 12 and 13, the vehicle carries (L_{i-1}, δ) units of goods from v_i to v_0, $(d_i - \delta, \delta)$ units of goods from v_0 to v_i, $(0, 0)$ units of goods from v_i to v_0, and

$(L_{i-1} + \lceil \frac{d_i - L_{i-1}}{\lambda-\delta} \rceil \cdot (\lambda - \delta) - d_i, \delta)$ units of goods from v_0 to v_i, and the cumulative cost is $a \cdot 4 \cdot l_i + b \cdot (L_{i-1} + \delta + d_i + 0 + L_{i-1} + \lceil \frac{d_i - L_{i-1}}{\lambda-\delta} \rceil \cdot (\lambda - \delta) - d_i + \delta) \cdot l_i = a \cdot 4 \cdot l_i + b \cdot (2L_{i-1} + 2\delta + \lceil \frac{d_i - L_{i-1}}{\lambda-\delta} \rceil \cdot (\lambda - \delta)) \cdot l_i$. Hence, the expected cumulative cost is

$$\begin{aligned}
&\int_{d_i-\delta}^{\min\{d_i,\lambda-\delta\}} \frac{a \cdot 2 \cdot l_i + b \cdot (2x - 2d_i + 2\delta + \lceil \frac{d_i - x}{\lambda-\delta} \rceil \cdot (\lambda - \delta)) \cdot l_i}{\lambda - \delta} dx \\
&+ \int_{0}^{d_i-\delta} \frac{a \cdot 4 \cdot l_i + b \cdot (2x + 2\delta + \lceil \frac{d_i - x}{\lambda-\delta} \rceil \cdot (\lambda - \delta)) \cdot l_i}{\lambda - \delta} dx \\
= &\int_{d_i-\delta}^{\lambda-\delta} \frac{a \cdot 2 \cdot l_i + b \cdot (2x - 2d_i + 2\delta + 1 \cdot (\lambda - \delta)) \cdot l_i}{\lambda - \delta} dx \\
&+ \int_{0}^{d_i+\delta-\lambda} \frac{a \cdot 4 \cdot l_i + b \cdot (2x + 2\delta + 2 \cdot (\lambda - \delta)) \cdot l_i}{\lambda - \delta} dx \\
&+ \int_{d_i+\delta-\lambda}^{d_i-\delta} \frac{a \cdot 4 \cdot l_i + b \cdot (2x + 2\delta + 1 \cdot (\lambda - \delta)) \cdot l_i}{\lambda - \delta} dx \\
= &\frac{a \cdot 2 \cdot (\lambda - d_i) \cdot l_i + b \cdot (d_i^2 + (\delta - 3\lambda) \cdot d_i + (2\lambda^2 - \lambda \cdot \delta))}{\lambda - \delta} \\
&+ \frac{a \cdot 4 \cdot (d_i + \delta - \lambda) \cdot l_i + b \cdot (d_i^2 + 2\delta \cdot d_i + \delta^2 - \lambda^2)}{\lambda - \delta} \\
&+ \frac{a \cdot 4 \cdot (\lambda - 2\delta) \cdot l_i + b \cdot ((2\lambda - 4\delta) \cdot d_i + \delta \cdot \lambda - 2\delta^2)}{\lambda - \delta} \\
= &\, a \cdot \frac{2d_i + 2\lambda - 4\delta}{\lambda - \delta} \cdot l_i + b \cdot \frac{2d_i^2 - (\lambda + \delta) \cdot d_i + \lambda^2 - \delta^2}{\lambda - \delta} \cdot l_i,
\end{aligned}$$

where the first equality follows from that $\lceil \frac{d_i - x}{\lambda-\delta} \rceil = 1$ if $d_i + \delta - \lambda \leq x \leq \lambda - \delta$ and $\lceil \frac{d_i - x}{\lambda-\delta} \rceil = 2$ if $0 \leq x < d_i + \delta - \lambda$ since in our setting $\delta \leq \lambda - \delta$ and in this case $\lambda - \delta < d_i \leq \lambda$. ◀

▶ **Theorem 9 (*).** *For Cu-VRPSD with any $\lambda \in (0, 1]$ and $\delta \in [0, \lambda/2]$, conditioned on $\chi = d$, $ALG.1(\lambda, \delta)$ generates a solution $\mathcal{T}$ with an expected cumulative cost of*

$$\frac{A + B}{\gamma \cdot \max\{\sigma,\ 1\} + 0.5} \cdot LB,$$

where $A \coloneqq \gamma \cdot (\alpha \cdot \sigma + \int_0^\delta \frac{x}{\lambda-\delta} dF(x) + \int_\delta^{\lambda-\delta} \frac{2x-\delta}{\lambda-\delta} dF(x) + \int_{\lambda-\delta}^\lambda \frac{x+\lambda-2\delta}{\lambda-\delta} dF(x) + \int_\lambda^1 1 dF(x))$, *and* $B \coloneqq (\frac{\lambda+\delta}{2} \cdot \alpha \cdot \sigma + \int_0^\delta \frac{(\lambda+\delta)x - x^2}{2(\lambda-\delta)} dF(x) + \int_\delta^{\lambda-\delta} \frac{x^2 + (\lambda-\delta)x}{2(\lambda-\delta)} dF(x) + \int_{\lambda-\delta}^\lambda \frac{2x^2 - (\lambda+\delta) \cdot x + \lambda^2 - \delta^2}{2(\lambda-\delta)} dF(x) + \int_\lambda^1 \frac{x}{2} dF(x))$.

3.1.2 The application

Next, we use $ALG.1(\lambda, \delta)$ to deign approximation algorithms for $a, b > 0$, i.e., $\gamma > 0$.

When the vehicle traveling along each edge of the TSP tour in $ALG.1(\lambda, \delta)$, it always carries at least δ units of goods in total, resulting in a large cumulative cost for the case where γ is small. Hence, intuitively, if γ is small, we simply set $\delta = 0$.

If we call $ALG.1(\lambda, \delta)$ with $\delta = 0$ and λ being a unique value, in the worst case we may have $\int_0^\lambda x^2 dF(x) = \int_0^\lambda \lambda \cdot x dF(x)$, i.e., almost all customers v_i with $l_i > 0$ and $d_i \leq \lambda$ have a demand of $d_i = \lambda$. Consequently, we may get $\int_0^{\theta \cdot \lambda} x dF(x) = 0$ for any fixed $\theta \in (0, 1)$, as will be shown in Lemma 10, and then $ALG.1(\theta \cdot \lambda, \delta)$ with $\delta = 0$ and some $\theta \in (0, 1)$ may generate a better solution. This suggests the approximation algorithm for Cu-VRPSD shown in Algorithm 2, denoted as $APPROX.1(\lambda, \theta, p)$.

Algorithm 2 An approximation algorithm for unsplittable Cu-VRPSD ($APPROX.1(\lambda, \theta, p)$).

Input: An instance of unsplittable Cu-VRPSD, and three parameters $\lambda \in (0, 1]$, $\theta \in (0, 1)$ and $p \in (0, 1)$.
Output: A feasible solution to Cu-VRPSD.
1: Call $ALG.1(\lambda, 0)$ with a probability of p and call $ALG.1(\theta \cdot \lambda, 0)$ with a probability of $1 - p$.

Then, our goal is to find (λ, θ, p) minimizing the approximation ratio of $APPROX.1(\lambda, \theta, p)$.

▶ **Lemma 10.** *For any $\theta \in (0, 1)$, we have $\int_0^{\theta\cdot\lambda} xdF(x) \leq \frac{1}{\lambda-\theta\cdot\lambda}(\int_0^{\lambda} \lambda \cdot xdF(x) - \int_0^{\lambda} x^2dF(x))$.*

Proof. By (1) and (2), we have $\int_0^{\lambda} x^2dF(x) = \int_0^{\theta\cdot\lambda} x^2dF(x) + \int_{\theta\cdot\lambda}^{\lambda} x^2dF(x) \leq \theta \cdot \lambda \cdot \int_0^{\theta\cdot\lambda} xdF(x) + \lambda \cdot \int_{\theta\cdot\lambda}^{\lambda} xdF(x) = \lambda \cdot \int_0^{\lambda} xdF(x) - (\lambda - \theta \cdot \lambda) \cdot \int_0^{\theta\cdot\lambda} xdF(x)$. Hence, we have $\int_0^{\theta\cdot\lambda} xdF(x) \leq \frac{1}{\lambda-\theta\cdot\lambda}(\int_0^{\lambda} \lambda \cdot xdF(x) - \int_0^{\lambda} x^2dF(x))$. ◀

▶ **Theorem 11.** *For unsplittable Cu-VRPSD, we can find (λ, θ, p) such that the approximation ratio of $APPROX.1(\lambda, \theta, p)$ is bounded by $10/3$ for any $\gamma \in (0, 0.375]$ and 3.456 for any $\gamma \in (0.375, 1.444]$.*

Proof. If we call $ALG.1(\lambda, \delta)$ with $\delta = 0$ and λ being a unique value, one may check that a good choice for λ is $\min\{1, 4\gamma/\alpha\}$. For the sake of analysis, we directly set $\lambda = \min\{1, 4\gamma/\alpha\}$.

If $\int_0^{\lambda} xdF(x) = 0$, we can get $\int_0^{\lambda} x^2dF(x) \leq \lambda \cdot \int_0^{\lambda} xdF(x) = 0$. Hence, we define $\mu := 0$ if $\int_0^{\lambda} xdF(x) = 0$, and $\mu := \frac{\int_0^{\lambda} x^2dF(x)}{\int_0^{\lambda} xdF(x)}$ otherwise.

By Lemma 4 and Theorem 9, the approximation ratio of $ALG.1(\lambda, 0)$ is at most

$$
\begin{aligned}
&\max_{\sigma\geq 0} \frac{\gamma \cdot \left(\alpha \cdot \sigma + \int_0^{\lambda} \frac{2x}{\lambda} dF(x) + \int_{\lambda}^{1} 1dF(x)\right) + \left(\frac{\lambda}{2} \cdot \alpha \cdot \sigma + \int_0^{\lambda} \frac{x^2+\lambda\cdot x}{2\lambda} dF(x) + \int_{\lambda}^{1} \frac{x}{2} dF(x)\right)}{\gamma \cdot \max\{\sigma,\ 1\} + 0.5} \\
&= \max_{\sigma\geq 0} \frac{\gamma \cdot \left(\alpha \cdot \sigma + \frac{2}{\lambda} \cdot \int_0^{\lambda} xdF(x) + \int_{\lambda}^{1} 1dF(x)\right) + \left(\frac{\lambda}{2} \cdot \alpha \cdot \sigma + \frac{1+\mu/\lambda}{2} \cdot \int_0^{\lambda} xdF(x) + \frac{1}{2} \cdot \int_{\lambda}^{1} xdF(x)\right)}{\gamma \cdot \max\{\sigma,\ 1\} + 0.5} \\
&\leq \max_{\sigma\geq 0} \frac{\gamma \cdot \left(\alpha \cdot \sigma + \frac{2}{\lambda} \cdot \int_0^{\lambda} xdF(x) + \frac{1}{\lambda} \cdot \int_{\lambda}^{1} xdF(x)\right) + \left(\frac{\lambda}{2} \cdot \alpha \cdot \sigma + \frac{1+\mu/\lambda}{2} \cdot \int_0^{\lambda} xdF(x) + \frac{1}{2} \cdot \int_{\lambda}^{1} xdF(x)\right)}{\gamma \cdot \max\{\sigma,\ 1\} + 0.5} \\
&= \max_{\sigma\geq 0} \frac{\gamma \cdot \left(\alpha \cdot \sigma + \frac{1}{\lambda} \cdot \int_0^{\lambda} xdF(x) + \frac{1}{\lambda}\right) + \left(\frac{\lambda}{2} \cdot \alpha \cdot \sigma + \frac{\mu/\lambda}{2} \cdot \int_0^{\lambda} xdF(x) + \frac{1}{2}\right)}{\gamma \cdot \max\{\sigma,\ 1\} + 0.5} \\
&\leq \max_{\sigma\geq 0} \frac{\gamma \cdot (\alpha \cdot \sigma + \frac{2}{\lambda}) + \left(\frac{\lambda}{2} \cdot \alpha \cdot \sigma + \frac{\mu/\lambda+1}{2}\right)}{\gamma \cdot \max\{\sigma,\ 1\} + 0.5},
\end{aligned}
$$

where the first equality follows from the definition of μ, the second equality from $\int_0^1 xdF(x) = 1$ by (1), the first inequality from $\int_{\lambda}^{1} 1dF(x) \leq \frac{1}{\lambda} \cdot \int_{\lambda}^{1} xdF(x)$ by (2), and the second inequality from $\int_0^{\lambda} xdF(x) \leq \int_0^1 xdF(x) = 1$.

Since $\int_0^{\lambda} xdF(x) \leq \int_0^1 xdF(x) = 1$, by Lemma 10, we have

$$
\begin{aligned}
\int_0^{\theta\cdot\lambda} xdF(x) &\leq \frac{1}{\lambda - \theta \cdot \lambda} \cdot \left(\int_0^{\lambda} \lambda \cdot xdF(x) - \int_0^{\lambda} x^2dF(x)\right) \\
&= \frac{\lambda - \mu}{\lambda - \theta \cdot \lambda} \cdot \int_0^{\lambda} xdF(x) \leq \frac{\lambda - \mu}{\lambda - \theta \cdot \lambda} \qquad (3)
\end{aligned}
$$

Similarly, the approximation ratio of $ALG.1(\theta \cdot \lambda, 0)$ is at most

$$\max_{\sigma\geq 0}\frac{\gamma\cdot\left(\alpha\cdot\sigma+\int_0^{\theta\cdot\lambda}\frac{2x}{\theta\cdot\lambda}dF(x)+\int_{\theta\cdot\lambda}^1 1dF(x)\right)+\left(\frac{\theta\cdot\lambda}{2}\cdot\alpha\cdot\sigma+\int_0^{\theta\cdot\lambda}\frac{x^2+\theta\cdot\lambda\cdot x}{2\cdot\theta\cdot\lambda}dF(x)+\int_{\theta\cdot\lambda}^1\frac{x}{2}dF(x)\right)}{\gamma\cdot\max\{\sigma,\ 1\}+0.5}$$
$$\leq\max_{\sigma\geq 0}\frac{\gamma\cdot\left(\alpha\cdot\sigma+\frac{2}{\theta\cdot\lambda}\cdot\int_0^{\theta\cdot\lambda}xdF(x)+\frac{1}{\theta\cdot\lambda}\cdot\int_{\theta\cdot\lambda}^1 xdF(x)\right)+\left(\frac{\theta\cdot\lambda}{2}\cdot\alpha\cdot\sigma+\int_0^{\theta\cdot\lambda}xdF(x)+\frac{1}{2}\cdot\int_{\theta\cdot\lambda}^1 xdF(x)\right)}{\gamma\cdot\max\{\sigma,\ 1\}+0.5}$$
$$=\max_{\sigma\geq 0}\frac{\gamma\cdot\left(\alpha\cdot\sigma+\frac{1}{\theta\cdot\lambda}\cdot\int_0^{\theta\cdot\lambda}xdF(x)+\frac{1}{\theta\cdot\lambda}\right)+\left(\frac{\theta\cdot\lambda}{2}\cdot\alpha\cdot\sigma+\frac{1}{2}\cdot\int_0^{\theta\cdot\lambda}xdF(x)+\frac{1}{2}\right)}{\gamma\cdot\max\{\sigma,\ 1\}+0.5}$$
$$\leq\max_{\sigma\geq 0}\frac{\gamma\cdot\left(\alpha\cdot\sigma+\frac{1}{\theta\cdot\lambda}\cdot\frac{\lambda-\mu}{\lambda-\theta\cdot\lambda}+\frac{1}{\theta\cdot\lambda}\right)+\left(\frac{\theta\cdot\lambda}{2}\cdot\alpha\cdot\sigma+\frac{1}{2}\cdot\frac{\lambda-\mu}{\lambda-\theta\cdot\lambda}+\frac{1}{2}\right)}{\gamma\cdot\max\{\sigma,\ 1\}+0.5}$$
$$=\max_{\sigma\geq 0}\frac{\gamma\cdot\left(\alpha\cdot\sigma+\frac{1}{\theta\cdot\lambda}\cdot\frac{2\lambda-\mu-\theta\cdot\lambda}{\lambda-\theta\cdot\lambda}\right)+\left(\frac{\theta\cdot\lambda}{2}\cdot\alpha\cdot\sigma+\frac{1}{2}\cdot\frac{2\lambda-\mu-\theta\cdot\lambda}{\lambda-\theta\cdot\lambda}\right)}{\gamma\cdot\max\{\sigma,\ 1\}+0.5},$$

where the first inequality follows from (2), the second inequality from (3), and the first equality from $\int_0^1 xdF(x)=1$ by (1).

Recall that in $APPROX.1(\lambda,\theta,p)$ we call $ALG.1(\lambda,0)$ (resp., $ALG.1(\theta\cdot\lambda,0)$) with a probability of p (resp., $1-p$). Hence, to erase the items related to μ in the numerators of the approximation ratios of $ALG.1(\lambda,0)$ and $ALG.1(\theta\cdot\lambda,0)$, we need to set p such that $p\cdot\frac{1}{2}\cdot\frac{1}{\lambda}+(1-p)\cdot(\gamma\cdot\frac{1}{\theta\cdot\lambda}\cdot\frac{-1}{\lambda-\theta\cdot\lambda}+\frac{1}{2}\cdot\frac{-1}{\lambda-\theta\cdot\lambda})=0$. Then, we can get $p=\frac{\frac{1}{2(\lambda-\theta\cdot\lambda)}+\frac{\gamma}{\theta\cdot\lambda(\lambda-\theta\cdot\lambda)}}{\frac{1}{2\lambda}+\frac{1}{2(\lambda-\theta\cdot\lambda)}+\frac{\gamma}{\theta\cdot\lambda(\lambda-\theta\cdot\lambda)}}$. Clearly, we have $p\in[0,1]$. Hence, the approximation ratio is $\max_{\sigma\geq 0}R(\sigma)$, where

$$R(\sigma):=\frac{\gamma\cdot\left(\alpha\cdot\sigma+p\cdot\frac{2}{\lambda}+(1-p)\cdot\frac{1}{\theta\cdot\lambda}\cdot\frac{2\lambda-\theta\cdot\lambda}{\lambda-\theta\cdot\lambda}\right)+\left(\frac{p\cdot\lambda+(1-p)\cdot\theta\cdot\lambda}{2}\cdot\alpha\cdot\sigma+p\cdot\frac{1}{2}+(1-p)\cdot\frac{1}{2}\cdot\frac{2\lambda-\theta\cdot\lambda}{\lambda-\theta\cdot\lambda}\right)}{\gamma\cdot\max\{\sigma,\ 1\}+0.5}.$$

It is easy to check $\max_{\sigma\geq 0}R(\sigma)=\max_{\sigma\geq 1}R(\sigma)$. Moreover, since the function $\frac{a'x+b'}{c'x+d'}$ with $x\geq 1$ and $a',b',c',d'>0$ attains the maximum value only if $x=1$ or $x=\infty$, we know that the approximation ratio is bounded by $\max\{R(1),\ R(\infty)\}$. Recall that $\lambda=\min\{1,4\gamma/\alpha\}$ and $p=\frac{\frac{1}{2(\lambda-\theta\cdot\lambda)}+\frac{\gamma}{\theta\cdot\lambda(\lambda-\theta\cdot\lambda)}}{\frac{1}{2\lambda}+\frac{1}{2(\lambda-\theta\cdot\lambda)}+\frac{\gamma}{\theta\cdot\lambda(\lambda-\theta\cdot\lambda)}}$. Assume $\alpha=1.5$, and then we have $\alpha/4=0.375$. By calculation, we have the following results.

- When $\gamma\in(0,0.375]$, setting $\theta=0.5$, we have $\max\{R(1),\ R(\infty)\}\equiv 10/3$;
- When $\gamma\in[0.375,1.444]$, setting $\theta=0.6677$, we have $\max\{R(1),\ R(\infty)\}\leq 3.456$.

The result for $\gamma\in(0,0.375]$ may be surprising. We give the details of its proof.

▷ Claim 12. When $\gamma\in(0,0.375]$, setting $\theta=0.5$, we have $\max\{R(1),\ R(\infty)\}\equiv 10/3$.

Proof. Note that $\lambda=\min\{1,4\gamma/\alpha\}=4\gamma/\alpha$ and $\alpha=1.5$. Setting $\theta=0.5$, we can get $p=\frac{\frac{1}{2(\lambda-\theta\cdot\lambda)}+\frac{\gamma}{\theta\cdot\lambda(\lambda-\theta\cdot\lambda)}}{\frac{1}{2\lambda}+\frac{1}{2(\lambda-\theta\cdot\lambda)}+\frac{\gamma}{\theta\cdot\lambda(\lambda-\theta\cdot\lambda)}}=\frac{5}{6}$. Hence, under $\sigma\geq 1$, we have

$$R(\sigma)=\frac{3/2\cdot\gamma\cdot\sigma+5/8+3/8+11/6\cdot\gamma\cdot\sigma+2/3}{\gamma\cdot\sigma+0.5}=\frac{10}{3}\cdot\frac{\gamma\cdot\sigma+0.5}{\gamma\cdot\sigma+0.5}=\frac{10}{3}.$$

Hence, we have $\max\{R(1),\ R(\infty)\}\equiv 10/3$. ◁

This finishes the proof. ◀

We mention that the approximation ratio of $APPROX.1$ may achieve $\alpha+2=3.5$ when $\gamma=\infty$. Hence, it can not improve the current best approximation algorithm for VRPSD [16]. Additionally, a more careful design than $APPROX.1$ could yield improved approximations; however, the optimal design remains unknown.

3.2 The second algorithm

In this section, we will introduce our second algorithm, denoted as $ALG.2(\lambda, \delta)$, which can be used to get a 3.456-approximation algorithm for Cu-VRPSD with any $\gamma \in (1.444, \infty)$, and an $(\alpha + 1.75 = 3.25)$-approximation algorithm for VRPSD. Here, we may require $\lambda \in (0, 1]$, $\delta \in (0, \lambda/2]$, and $1/\delta \in \mathbb{N}$.

$ALG.2(\lambda, \delta)$ is based on $ALG.1(\lambda, \delta)$. The vehicle will skip customers v_i with $d_i > \lambda$ when it travels along the TSP tour in $ALG.1(\lambda, \delta)$, and then satisfy each of them using a single tour at last. In $ALG.2(\lambda, \delta)$, the main difference is that the vehicle will skip customers v_i with $d_i > \delta$, and at last use the better method from either satisfying each of them using a single tour or solving the weighted $(1 - \delta)/\delta$-set cover problem as shown below.

Given a *feasible* set of unsatisfied customers S such that the total demand of all customers in S is at most 1, we know that the number of customers $|S|$ is at most $(1 - \delta)/\delta$ since each unsatisfied customer v_i has a demand of $d_i > \delta$. Then, we can optimally compute a tour with a minimum cumulative cost $Cu(S)$ for all customers in S in $O(|S|!)$ time. There are at most $n^{O(1/\delta)}$ number of feasible sets since $|S| \leq (1 - \delta)/\delta$. Therefore, to satisfy customers v_i with $d_i > \delta$, we can get an instance of weighted $(1 - \delta)/\delta$-set cover by taking each unsatisfied customer as an element, and each feasible set S of unsatisfied customers as a set with a weight of $Cu(S)$ in polynomial time. By calling a ρ-approximation algorithm for weighted $(1 - \delta)/\delta$-set cover [15], we can get a set of tours satisfying all customers v_i with $d_i > \delta$.

According to the two methods, there are two set of tours $\mathcal{T}_1$ and $\mathcal{T}_2$, and their cumulative cost can be computed in polynomial time. Hence, we route the vehicle according to the tours in $\mathcal{T}'$, where $\mathcal{T}' \coloneqq \mathcal{T}_1$ if $Cu(\mathcal{T}_1) \leq Cu(\mathcal{T}_2)$ and $\mathcal{T}' \coloneqq \mathcal{T}_2$ otherwise.

The details of $ALG.2(\lambda, \delta)$ is shown in Algorithm 3.

▶ **Theorem 13 (*).** *For Cu-VRPSD with any* $\lambda \in (0, 1]$, $\delta \in (0, \lambda/2]$, *and* $1/\delta \in \mathbb{N}$, *conditioned on* $\chi = d$, $ALG.2(\lambda, \delta)$ *outputs a solution* $\mathcal{T}$ *with an expected cumulative cost of*

$$\frac{\gamma \cdot \left(\alpha \cdot \sigma + \int_0^\delta \frac{x}{\lambda - \delta} dF(x)\right) + \left(\frac{\lambda + \delta}{2} \cdot \alpha \cdot \sigma + \int_0^\delta \frac{(\lambda + \delta)x - x^2}{2(\lambda - \delta)} dF(x)\right)}{\gamma \cdot \max\{\sigma,\ 1\} + 0.5} \cdot LB + Cu(\mathcal{T}'),$$

where

$$Cu(\mathcal{T}') \leq \min\left\{\frac{\int_\delta^1 \frac{2\gamma + x}{2} dF(x)}{\gamma \cdot \max\{\sigma,\ 1\} + 0.5} \cdot LB,\ \ \rho \cdot Cu(\mathcal{T}^*)\right\}.$$

3.2.1 The applications

Our goal is to obtain a 3.456-approximation algorithm for Cu-VRPSD with any $\gamma \in (1.444, \infty)$. As a byproduct, we will also get a 3.25-approximation algorithm for VRPSD.

Since in $APPROX.1(\lambda, \theta, p)$ $ALG.1(\lambda, \delta)$ sets $\lambda = 1$ for any $\gamma > 0.375$, we also set $\lambda = 1$ in $ALG.2(\lambda, \delta)$ for the sake of analysis. Moreover, since weighted 2-set cover [15] can be solved optimally in polynomial time, i.e., $\rho = 1$ when $\delta = 1/3$, we set $\delta = 1/3$ in $ALG.2(\lambda, \delta)$.

According to Theorems 9 and 13, we will show that $ALG.2(\lambda, \delta)$ can be used to make a trade-off with $ALG.1(\lambda, \delta)$. We use the approximation algorithm for Cu-VRPSD shown in Algorithm 4, denoted as $APPROX.2$.

▶ **Theorem 14 (*).** *For unsplittable Cu-VRPSD,* $APPROX.2$ *is a randomized* 3.456-*approximation algorithm for any* $\gamma \in (1.444, \infty)$. *Moreover, for unsplittable VRPSD,* $APPROX.2$ *is a randomized* 3.25-*approximation algorithm.*

Combining the results in Lemma 5, Theorems 11 and 14, we get the following result.

Algorithm 3 An algorithm for unsplittable Cu-VRPSD ($ALG.2(\lambda, \delta)$).

Input: An instance of unsplittable Cu-VRPSD, and two parameters $\lambda \in (0, 1]$, $\delta \in (0, \lambda/2]$, and $1/\delta \in \mathbb{N}$.
Output: A feasible solution $\mathcal{T}$ to unsplittable Cu-VRPSD.

Obtain an α-approximate TSP tour $T^* = v_0 v_1 v_2 \ldots v_n v_0$, as Step 1 in $ALG.1(\lambda, \delta)$.
Load the vehicle with $S_0 := (L_0, \delta)$ units of goods, including L_0 demand of normal goods and δ demand of backup goods, where $L_0 \sim U[0, \lambda - \delta)$.
Initialize $i := 1$ and $V^* := \emptyset$.
while $i \leq n$ **do**
    Go to customer v_i;
    if $\delta < d_i \leq 1$ **then**
        Record v_i's demand, and let $V^* := V^* \cup \{v_i\}$ and $S_i := (L_i, \delta)$, where $L_i := L_{i-1}$;
    else if $d_i \leq L_{i-1}$ **then**
        Deliver $(d_i, 0)$ units of goods to v_i, and update $S_i := (L_i, \delta)$, where $L_i := L_{i-1} + \lceil \frac{d_i - L_{i-1}}{\lambda - \delta} \rceil \cdot (\lambda - \delta) - d_i = L_{i-1} - d_i$;
    else ▷ Since $d_i \leq \delta$, we must have $L_{i-1} < d_i \leq L_{i-1} + \delta$
        Deliver $(L_{i-1}, d_i - L_{i-1})$ units of goods to v_i, goes to the depot, load the vehicle with $(L_{i-1} + \lceil \frac{d_i - L_{i-1}}{\lambda - \delta} \rceil \cdot (\lambda - \delta) - d_i, d_i - L_{i-1})$ units of goods, and update $S_i := (L_i, \delta)$, where $L_i := L_{i-1} + \lceil \frac{d_i - L_{i-1}}{\lambda - \delta} \rceil \cdot (\lambda - \delta) - d_i = L_{i-1} + (\lambda - \delta) - d_i$;
    end if
    $i := i + 1$.
end while
Go to the depot.
Consider a set of tours $\mathcal{T}_1$ by obtaining a single tour as in Step 22 for each $v_i \in V^*$.
Consider a set of tours $\mathcal{T}_2$ by calling a ρ-approximation algorithm for weighted $\frac{1-\delta}{\delta}$-set cover [15], where the instance is constructed as follows:
1. Obtain all possible feasible sets S of customers in V^* such that the total demand of all customers in S is at most 1;
2. For each feasible set S, compute a tour with a minimum cumulative cost $Cu(S)$ for all customers in S;
3. Get an instance of weighted $\frac{1-\delta}{\delta}$-set cover by taking each customer in V^* as an element, and each feasible set S as a weighted set with a weight of $Cu(S)$.

Let $\mathcal{T}' := \mathcal{T}_1$ if $Cu(\mathcal{T}_1) \leq Cu(\mathcal{T}_2)$ and $\mathcal{T}' := \mathcal{T}_2$ otherwise.
Route the vehicle according to the tours in $\mathcal{T}'$.

Algorithm 4 An approximation algorithm for unsplittable Cu-VRPSD ($APPROX.2$).

Input: An instance of unsplittable Cu-VRPSD.
Output: A feasible solution to Cu-VRPSD.

Call $ALG.1(1, 1/3)$ with a probability of 0.5 and call $ALG.2(1, 1/3)$ with a probability of 0.5.

▶ **Corollary 15.** *There is a randomized* 3.456*-approximation for unsplittable Cu-VRPSD.*

▶ Remark 16. We believe that our analysis is not tight. One one hand, it would be interesting to sharpen our analysis to get a better result; on the other hand, we may use $ALG.1$ and $ALG.2$ to design better approximation algorithms, e.g., with a probability of p_γ to run $ALG.1$, and of $(1 - p_\gamma)$ to run $ALG.2$, where p_γ is a function related to γ. Moreover, when running $ALG.1$ or $ALG.2$, the parameters λ and δ may follow a distribution related to γ.

4 Two Algorithms for Cu-VRP

In this section, we give a 3.194-approximation algorithm for Cu-VRP.

4.1 The first algorithm

Based on the well-known randomized rounding method for weighted k-set cover, we propose a 3.194-approximation algorithm, denoted as $ALG.3(\lambda, \delta)$, for Cu-VRP with any $\gamma > \gamma_0$, where $\gamma_0 > 0$ is any fixed constant.

Recall that $ALG.2(\lambda, \delta)$ first satisfies customers v_i with $d_i \leq \delta$ by traveling along the TSP tour and then customers v_i with $d_i > \delta$ by possibly solving weighted $\frac{1-\delta}{\delta}$-set cover. However, it may only be used for $\delta = 1/3$ since the best-known approximation ratio of weighted 3-set cover is about 1.79 [15], which is already too large.

In $ALG.3(\lambda, \delta)$, since the demands of customers are known in advance for Cu-VRP, we first try to satisfy customers in $V^* := \{v_i \in V' \mid d_i > \delta\}$ by solving weighted $\frac{1-\delta}{\delta}$-set cover using the randomized rounding method. Due to the randomness, some customers in V^* may still be unsatisfied. Then, we satisfy all remaining customers by calling $ALG.1(\lambda, \delta)$. This method was used to get an $(\alpha + 1 + \ln 2 + \varepsilon)$-approximation algorithm with any constant $\varepsilon > 0$ for unsplittable VRP [8]. The details are shown as follows.

To get an instance of weighted $\frac{1-\delta}{\delta}$-set cover, we use the method in Step 17 of $ALG.2(\lambda, \delta)$. Now, we have obtained a set of feasible sets $\mathcal{S}$, and each $S \in \mathcal{S}$ has a weight of $Cu(S)$. Then, we get the linear relaxation of weighted set cover as shown in (4), and it can be solved in $n^{O(1/\delta)}$ since $|\mathcal{S}| = n^{O(1/\delta)}$. In the randomized rounding method, we select each $S \in \mathcal{S}$ with a probability of $\min\{\ln 2 \cdot x_S, 1\}$. Denote the set of selected sets by $\mathcal{S}'$, which corresponds to a set of tours $\mathcal{T}'$ satisfying a subset of customers $V^{**} \subseteq V^*$. Note that $Cu(\mathcal{T}') \leq Cu(\mathcal{S}')$ since we may perform shortcutting to ensure that each customer appears in only one tour and it does not increase the cumulative cost by the triangle inequality. At last, we call $ALG.1(\lambda, \delta)$ to obtain a set of tours $\mathcal{T}''$ to satisfy the left customers in $V' \setminus V^{**}$. Due to the stochastic demands in Cu-VRPSD the load of the vehicle may be greater than the delivered units of goods in each tour of $\mathcal{T}''$. In Cu-VRP, we can optimize the tours in $\mathcal{T}''$ so that the load equals the delivered units of goods. Moreover, for each tour $T = v_0 v_{i_1} \ldots v_{i_T} v_0 \in \mathcal{T}''$, we consider another tour with the opposite direction, i.e., $v_0 v_{i_T} \ldots v_{i_1} v_0$, and choose the better one into our final solution.

$$\begin{aligned} \text{minimize} \quad & \sum_{S \in \mathcal{S}} Cu(S) \cdot x_S \\ \text{subject to} \quad & \sum_{S \in \mathcal{S} : v \in S} x_S \geq 1, \quad \forall\, v \in V^*, \\ & x_S \geq 0, \quad \forall\, S \in \mathcal{S}. \end{aligned} \tag{4}$$

The details of $ALG.3(\lambda, \delta)$ is shown in Algorithm 5.

▶ **Theorem 17 (*).** *For Cu-VRP with any $\lambda \in (0, 1]$, $\delta \in (0, \lambda/2]$, and $1/\delta \in \mathbb{N}$, $ALG.3(\lambda, \delta)$ generates a solution $\mathcal{T}$ with an expected cumulative cost of*

$$\ln 2 \cdot Cu(\mathcal{T}^*) + \frac{\gamma \cdot \left(\alpha \cdot \sigma + \frac{1}{\lambda - \delta}\right) + \frac{\lambda}{2} \cdot \left(\alpha \cdot \sigma + \frac{1}{\lambda - \delta}\right)}{\gamma \cdot \max\{\sigma,\ 1\} + 0.5} \cdot LB.$$

▶ **Theorem 18 (*).** *For unsplittable Cu-VRP with any constants $\gamma_0 > 0$ and $\varepsilon > 0$, there is a randomized $(\alpha + 1 + \ln 2 + \varepsilon < 3.194)$-approximation algorithm for $\gamma > \gamma_0$.*

4.2 The second algorithm

In this section, we propose a 3.163-approximation algorithm for Cu-VRP with $\gamma \in (0, 0.428]$, denoted as $ALG.4(\lambda)$. Combing with Lemma 5 and Theorem 18, $ALG.4(\lambda)$ implies a 3.194-approximation algorithm for Cu-VRP.

Algorithm 5 An algorithm for unsplittable Cu-VRPSD ($ALG.3(\lambda, \delta)$).

Input: An instance of unsplittable Cu-VRPSD, and two parameters $\lambda \in (0, 1]$, $\delta \in (0, \lambda/2]$, and $1/\delta \in \mathbb{N}$.
Output: A feasible solution $\mathcal{T}$ to unsplittable Cu-VRPSD.

1: Get an instance $(V^*, \mathcal{S})$ of weighted $\frac{1-\delta}{\delta}$-set cover using Step 17 in $ALG.2(\lambda, \delta)$.
2: Solve the linear program of weighted set cover in (4).
3: Select each $S \in \mathcal{S}$ with a probability of $\min\{\ln 2 \cdot x_S, 1\}$. Denote the set of selected sets by $\mathcal{S}'$, corresponding tours by $\mathcal{T}'$, and satisfied customers by V^{**}.
4: Call $ALG.1(\lambda, \delta)$ to obtain a set of tours $\mathcal{T}''$ to satisfy the customers in $V' \setminus V^{**}$.
5: For each tour in $\mathcal{T}''$, ensure the load of the vehicle is the delivered units of goods, obtain another tour with the opposite direction, and choose the better one into $\mathcal{T}'''$.
6: Return $\mathcal{T}' \cup \mathcal{T}'''$.

Algorithm 6 An algorithm for unsplittable Cu-VRPSD ($ALG.4(\lambda)$).

Input: An instance of unsplittable Cu-VRPSD, and two parameters $\lambda \in (0, 1]$, $\delta \in (0, \lambda/2]$, and $1/\delta \in \mathbb{N}$.
Output: A feasible solution $\mathcal{T}$ to unsplittable Cu-VRPSD.

1: Call $ALG.1(\lambda, 0)$ to obtain a set of tours $\mathcal{T}'$ to satisfy all customers.
2: For each tour in $\mathcal{T}'$, ensure the load of the vehicle is the delivered units of goods, obtain another tour with the opposite direction, and choose the better one into $\mathcal{T}$.
3: Return $\mathcal{T}$.

Algorithm 7 An approximation algorithm for unsplittable Cu-VRPSD ($APPROX.4(\lambda, \theta, p)$).

Input: An instance of unsplittable Cu-VRP, and three parameters $\lambda \in (0, 1]$, $\theta \in (0, 1)$ and $p \in (0, 1)$.
Output: A feasible solution to Cu-VRPSD.

1: Call $ALG.4(\lambda)$ with a probability of p and call $ALG.4(\theta \cdot \lambda)$ with a probability of $1 - p$.

In $ALG.4(\lambda)$, we call $ALG.1(\lambda, 0)$ to obtain a set of tours $\mathcal{T}'$ to satisfy all customers, and then we optimize each tour in $\mathcal{T}'$ as Step 5 in $ALG.3$.

▶ **Theorem 19 (*).** *For Cu-VRP with any $\lambda \in (0, 1]$, $ALG.4(\lambda)$ generates a solution $\mathcal{T}$ with an expected cumulative cost of*

$$\frac{\gamma \cdot \left(\alpha \cdot \sigma + \int_0^\lambda \frac{2x}{\lambda} dF(x) + \int_\lambda^1 1 dF(x)\right) + \left(\frac{\lambda}{2} \cdot \alpha \cdot \sigma + \int_0^\lambda \frac{x^2/2 + \lambda \cdot x}{2\lambda} dF(x) + \int_\lambda^1 \frac{x}{2} dF(x)\right)}{\gamma \cdot \max\{\sigma,\ 1\} + 0.5} \cdot LB.$$

Similarly, we use $ALG.4(\lambda)$ to design an algorithm for Cu-VRP shown in Algorithm 7.

▶ **Theorem 20 (*).** *For unsplittable Cu-VRP, we can find (λ, θ, p) such that the approximation ratio of $APPROX.4(\lambda, \theta, p)$ is bounded by 3.163 for any $\gamma \in (0, 0.428]$.*

5 Conclusion

By using the idea of skipping customers with large demands during the TSP tour and satisfying them later, combined with careful analysis, we can improve the approximation ratio for Cu-VRPSD, VRPSD, and Cu-VRP. Whether this idea is also useful in designing practical algorithms for these problems is worthy of further study.

References

1 Kemal Altinkemer and Bezalel Gavish. Heuristics for unequal weight delivery problems with a fixed error guarantee. *Operations Research Letters*, 6(4):149–158, 1987.

2 Dimitris J. Bertsimas. A vehicle routing problem with stochastic demand. *Oper. Res.*, 40(3):574–585, 1992. doi:10.1287/opre.40.3.574.

3 Dimitris J. Bertsimas and David Simchi-Levi. A new generation of vehicle routing research: Robust algorithms, addressing uncertainty. *Oper. Res.*, 44(2):286–304, 1996. doi:10.1287/opre.44.2.286.

4 Jannis Blauth, Vera Traub, and Jens Vygen. Improving the approximation ratio for capacitated vehicle routing. *Math. Program.*, 197(2):451–497, 2023. doi:10.1007/s10107-022-01841-4.

5 Nicos Christofides. Worst-case analysis of a new heuristic for the travelling salesman problem. Technical report, Carnegie-Mellon University, 1976.

6 Karina Corona-Gutiérrez, Samuel Nucamendi-Guillén, and Eduardo Lalla-Ruiz. Vehicle routing with cumulative objectives: A state of the art and analysis. *Comput. Ind. Eng.*, 169:108054, 2022. doi:10.1016/j.cie.2022.108054.

7 George B Dantzig and John H Ramser. The truck dispatching problem. *Management Science*, 6(1):80–91, 1959.

8 Zachary Friggstad, Ramin Mousavi, Mirmahdi Rahgoshay, and Mohammad R. Salavatipour. Improved approximations for capacitated vehicle routing with unsplittable client demands. In Karen I. Aardal and Laura Sanità, editors, *Integer Programming and Combinatorial Optimization - 23rd International Conference, IPCO 2022, Eindhoven, The Netherlands, June 27-29, 2022, Proceedings*, volume 13265 of *Lecture Notes in Computer Science*, pages 251–261. Springer, 2022. doi:10.1007/978-3-031-06901-7_19.

9 Ricardo Fukasawa, Qie He, and Yongjia Song. A branch-cut-and-price algorithm for the energy minimization vehicle routing problem. *Transp. Sci.*, 50(1):23–34, 2016. doi:10.1287/trsc.2015.0593.

10 Daya Ram Gaur, Apurva Mudgal, and Rishi Ranjan Singh. Routing vehicles to minimize fuel consumption. *Oper. Res. Lett.*, 41(6):576–580, 2013. doi:10.1016/j.orl.2013.07.007.

11 Daya Ram Gaur, Apurva Mudgal, and Rishi Ranjan Singh. Approximation algorithms for cumulative VRP with stochastic demands. In Sathish Govindarajan and Anil Maheshwari, editors, *Algorithms and Discrete Applied Mathematics - Second International Conference, CALDAM 2016, Thiruvananthapuram, India, February 18-20, 2016, Proceedings*, volume 9602 of *Lecture Notes in Computer Science*, pages 176–189. Springer, 2016. doi:10.1007/978-3-319-29221-2_15.

12 Daya Ram Gaur, Apurva Mudgal, and Rishi Ranjan Singh. Improved approximation algorithms for cumulative VRP with stochastic demands. *Discret. Appl. Math.*, 280:133–143, 2020. doi:10.1016/j.dam.2018.01.012.

13 Daya Ram Gaur and Rishi Ranjan Singh. A heuristic for cumulative vehicle routing using column generation. *Discret. Appl. Math.*, 228:140–157, 2017. doi:10.1016/j.dam.2016.05.030.

14 Michel Gendreau, Gilbert Laporte, and René Séguin. Stochastic vehicle routing. *Eur. J. Oper. Res.*, 88(1):3–12, 1996.

15 Anupam Gupta, Euiwoong Lee, and Jason Li. A local search-based approach for set covering. In Telikepalli Kavitha and Kurt Mehlhorn, editors, *2023 Symposium on Simplicity in Algorithms, SOSA 2023, Florence, Italy, January 23-25, 2023*, pages 1–11. SIAM, 2023. doi:10.1137/1.9781611977585.ch1.

16 Anupam Gupta, Viswanath Nagarajan, and R. Ravi. Technical note - approximation algorithms for VRP with stochastic demands. *Oper. Res.*, 60(1):123–127, 2012. doi:10.1287/opre.1110.0967.

17 Mordecai Haimovich and Alexander H. G. Rinnooy Kan. Bounds and heuristics for capacitated routing problems. *Math. Oper. Res.*, 10(4):527–542, 1985. doi:10.1287/moor.10.4.527.

18 İmdat Kara, Bahar Yetiş Kara, and M Kadri Yetis. Energy minimizing vehicle routing problem. In *Combinatorial Optimization and Applications: First International Conference, COCOA 2007*, pages 62–71. Springer, 2007.

19 İmdat Kara, Bahar Yetiş Kara, and M Kadri Yetis. *Cumulative vehicle routing problems*. IntechOpen Rijeka, HR, 2008.

20 Anna R. Karlin, Nathan Klein, and Shayan Oveis Gharan. A (slightly) improved approximation algorithm for metric TSP. In Samir Khuller and Virginia Vassilevska Williams, editors, *STOC '21: 53rd Annual ACM SIGACT Symposium on Theory of Computing, Virtual Event, Italy, June 21-25, 2021*, pages 32–45. ACM, 2021. doi:10.1145/3406325.3451009.

21 Anna R. Karlin, Nathan Klein, and Shayan Oveis Gharan. A deterministic better-than-3/2 approximation algorithm for metric TSP. In Alberto Del Pia and Volker Kaibel, editors, *Integer Programming and Combinatorial Optimization - 24th International Conference, IPCO 2023, Madison, WI, USA, June 21-23, 2023, Proceedings*, volume 13904 of *Lecture Notes in Computer Science*, pages 261–274. Springer, 2023. doi:10.1007/978-3-031-32726-1_19.

22 Mauro Henrique Mulati, Ricardo Fukasawa, and Flávio Keidi Miyazawa. The arc-item-load and related formulations for the cumulative vehicle routing problem. *Discret. Optim.*, 45:100710, 2022. doi:10.1016/j.disopt.2022.100710.

23 Bahri Sahin, Huseyin Yilmaz, Yasin Ust, Ali Fuat Guneri, and Bahadir Gülsün. An approach for analysing transportation costs and a case study. *Eur. J. Oper. Res.*, 193(1):1–11, 2009. doi:10.1016/j.ejor.2007.10.030.

24 Anatolii Ivanovich Serdyukov. Some extremal bypasses in graphs. *Upravlyaemye Sistemy*, 17:76–79, 1978.

25 Xinyu Wang, Tsan-Ming Choi, Haikuo Liu, and Xiaohang Yue. A novel hybrid ant colony optimization algorithm for emergency transportation problems during post-disaster scenarios. *IEEE Trans. Syst. Man Cybern. Syst.*, 48(4):545–556, 2016.

26 Yuanxiao Wu and Xiwen Lu. Capacitated vehicle routing problem on line with unsplittable demands. *J. Comb. Optim.*, 44(3):1953–1963, 2022. doi:10.1007/s10878-020-00565-5.

27 Yiyong Xiao, Qiuhong Zhao, Ikou Kaku, and Yuchun Xu. Development of a fuel consumption optimization model for the capacitated vehicle routing problem. *Comput. Oper. Res.*, 39(7):1419–1431, 2012. doi:10.1016/j.cor.2011.08.013.