



**HAL**  
open science

# The electric vehicle routing problem with capacitated charging stations

Aurélien Froger, Jorge E. Mendoza, Ola Jabali, Gilbert Laporte

► **To cite this version:**

Aurélien Froger, Jorge E. Mendoza, Ola Jabali, Gilbert Laporte. The electric vehicle routing problem with capacitated charging stations. 2019. hal-02386167v1

**HAL Id: hal-02386167**

**<https://hal.science/hal-02386167v1>**

Preprint submitted on 29 Nov 2019 (v1), last revised 31 Dec 2021 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The electric vehicle routing problem with capacitated charging stations

Aurélien Froger<sup>1</sup>    Jorge E. Mendoza<sup>2</sup>    Ola Jabali<sup>3</sup>    Gilbert Laporte<sup>2</sup>

<sup>1</sup> Université de Bordeaux, UMR CNRS 5251 – Inria Bordeaux Sud-Ouest

<sup>2</sup> HEC Montréal, Montréal, Canada H3T 2A7

<sup>3</sup> Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milan, Italy

*Submitted for publication in November 2019*

## Abstract

Much of the existing research on electric vehicle routing problems (E-VRPs) assumes that the charging stations (CSs) can simultaneously charge an unlimited number of electric vehicles, but this is not the case. In this research, we investigate how to model and solve E-VRPs taking into account these capacity restrictions. In particular, we study an E-VRP with non-linear charging functions, multiple charging technologies, en route charging, and variable charging quantities, while explicitly accounting for the capacity of CSs expressed in the number of chargers. We refer to this problem as the E-VRP with non-linear charging functions and capacitated stations (E-VRP-NL-C). This problem advances the E-VRP literature by considering the scheduling of charging operations at each CS. We first introduce two mixed integer linear programming formulations showing how CS capacity constraints can be incorporated into E-VRP models. We then introduce an algorithmic framework to the E-VRP-NL-C, that iterates between two main components: *a route generator* and *a solution assembler*. The route generator uses an iterated local search algorithm to build a pool of high-quality routes. The solution assembler applies a branch-and-cut algorithm to select a subset of routes from the pool. We report on computational experiments comparing four different assembly strategies on a large and diverse set of instances. Our results show that our algorithm deals with the CS capacity constraints effectively. Furthermore, considering the well-known uncapacitated version of the E-VRP-NL-C, our solution method identifies new best-known solutions for 80 out of 120 instances.

**Keywords**— Electric vehicle routing; non-linear charging function; synchronization constraints; mixed integer linear programming; matheuristic; iterated local search; branch-and-cut

## 1 Introduction

In recent years, competitive prices and technological advances have made electric vehicles (EVs) an attractive alternative to internal combustion engine-powered vehicles for logistics operations (Juan et al., 2016; Pelletier et al., 2016). Consequently, the operations research community has started paying attention to the study of electric vehicle routing problems (E-VRPs), which consist of designing routes to serve a set of customers using a fleet of EVs. Due to their relatively short driving range, EVs may need to detour to charging stations (CSs) to replenish their battery, especially in the context of mid-haul or long-haul routing (Schiffer et al., 2018; Villegas et al., 2018). Therefore, key decisions in E-VRPs concern not only the sequence in which the customers are served, but also where and by how much to charge the vehicles.

One of the main modeling elements in E-VRPs is the charging process of batteries. Some studies assume that EVs are fully recharged whenever they detour to a CS. In the green vehicle routing problem (G-VRP) tackled by Erdoğan and Miller-Hooks (2012), Koç and Karaoglan (2016), Montoya et al. (2016), and Bartolini and Andelmin (2017), the charging time is assumed to be constant, while in the works of Schneider et al. (2014), Hiermann et al. (2016), Keskin

and Çatay (2016), and Desaulniers et al. (2016) the charging time linearly depends on the state of charge (SoC) of the EV upon arrival at the CS. A full charging policy may be too restrictive, especially when considering en route charging. As Montoya et al. (2017) pointed out, this policy may lead to unnecessary out-of-the-depot charging, which translates into expensive driver idling time and overpriced energy purchases. To overcome this drawback, several researchers have studied problem variants in which the charging time is a decision variable (Desaulniers et al. (2016), Froger et al. (2019), Felipe et al. (2014), and Montoya et al. (2017)). The latter three papers examine the case in which the CSs may have different charging speeds. In reality, however, the battery charging process follows a non-linear function with respect to time (Pelletier et al. (2017)). To account for this, Montoya et al. (2017), Koç et al. (2018), and Froger et al. (2019) have modeled the charging process through concave piecewise linear functions. In the resulting problem, known as the electric vehicle routing problem with non-linear charging function (E-VRP-NL), the charging times are estimated more precisely than in the variants that assume a linear charging process.

All papers just reviewed assume that charging stations are always available, meaning that EVs do not need to wait to start charging. This implicitly means that the E-VRP research has mainly focused on problem variants where the charging infrastructure is privately owned by the route planner. In some contexts, this is a plausible assumption, since large companies may decide to invest in their own infrastructure to avoid dealing with the uncertainty in the availability of public CSs (Villegas et al., 2018). To the best of our knowledge, there exist only on a handful of references dealing with public charging stations for E-VRPs (see Sweda et al. (2017) and Kullman et al. (2018) for further details). Keskin et al. (2019) considered deterministic time-dependent queueing times at the stations.

Another common assumption in most, if not all, E-VRP publications is that CSs are always readily available to charge a vehicle when it arrives. Thus, implicitly assuming that CSs are uncapacitated and can simultaneously charge an unlimited number of EVs. In practice, however, each CS has a fixed and often small number of chargers. The intuition behind neglecting the CS capacity constraints is that accounting for the detour and charging times (or costs) while planning the routes is enough to capture the impact of the charging decisions on the cost and feasibility of solutions. Nonetheless, the long charging times (from tens of minutes to several hours) and the small number of chargers typically available at private CSs may generate congestion. To illustrate this point, we ran a feasibility test on the 120 best-known solutions (BKSs) for the E-VRP-NL reported in Montoya et al. (2017) limiting the number of chargers per CS to one, two, three, and four. According to our results, 55 of the BKS become infeasible if there is only one charger per CS. This figure drops to 23 and three for the cases with two and three chargers. The only scenario where all BKSs are feasible is when CSs have four chargers (see Appendix A for details). Intuitively, one may think that by merely shifting the starting time of the charging operations marginally, the feasibility problem will be solved. However, our experiments show that this is not only unnecessarily expensive, but it may also be infeasible. Another option to mitigate the effects of congestion is to increase the number of chargers installed at each CS, but this may not be a viable solution in practice. Indeed, Gnann et al. (2018) predict that in 2020, the purchase and installation costs of a fast charging point (i.e., power rates above 22 kW) will be around 40,000€ and its annual operation cost will reach 4,000€. Thus, if a company decides to invest in out-of-the-depot charging infrastructure, there is little chance that it will decide to install more than a couple of chargers at each CS. In conclusion, we argue that in most practical situations, the congestion at the CSs should be taken into account when planning the routes by explicitly modeling the CS capacity constraints.

We are aware of only one study taking into account CS capacity constraints when optimizing routing decisions. Bruglieri et al. (2019) explicitly limited the number of vehicles simultaneously refueling at each alternative fuelling station according to its number of fueling pumps. Some researchers have studied related but different problems. For instance, Sassi and Oulamara (2014) and Pelletier et al. (2018) considered the problem of scheduling charging operations at the depot, assuming that the routes are given as an input. Both papers considered not only constraints on the number of available chargers, but also electricity grid constraints limiting the amount of power that can be drawn from the electric grid at any given time. The latter is typically a binding constraint for central depots with a large number of available chargers, but is usually not a concern for CSs since they are designed to operate at full capacity.

In this paper we introduce the E-VRP-NL with capacitated CSs (E-VRP-NL-C). The problem extends classical E-VRPs to account for the fixed number of chargers available at each CS. The E-VRP-NL-C is a complex combined routing-scheduling problem belonging to the family of VRPs with synchronization constraints. More specifically, according to the taxonomy introduced by Drexl (2012), the E-VRP-NL-C belongs to the class of VRPs with *resource synchronization*, where vehicles compete to access scarce resources (i.e., the chargers at every CS). The E-VRP-NL-C shares similarities

with problems where vehicles need to wait while the loading equipment is busy. Examples of problems in this class include the log truck scheduling problem (El Hachemi et al., 2013; Rix et al., 2015) and routing and scheduling problems arising in public works (Grimault et al., 2017). The E-VRP-NL-C also relates to VRPs with inter-tour resource constraints. In these problems, the scarce resources are located only at the terminal node of the routes (i.e., the depot). An example of a problem in this category is the VRP introduced by Hemsch and Irnich (2008), where there is a limited number of ramps at the depot, and therefore only a fixed number of vehicles can be served simultaneously. The problem that is most closely related to the E-VRP-NL-C is the VRP with location congestion introduced by Lam and Van Hentenryck (2016). In this problem, each customer has a given number of requests and a limited number of resources (e.g., conveyors). When a vehicle arrives at a customer location, it must wait until at least one resource becomes available to handle the shipping. However, there exist two fundamental differences between their problem and our E-VRP-NL-C. The first is that in our problem, visiting nodes with limited resources (the CSs) is needed for feasibility reasons. As a consequence, not only the number, but also the location and duration of the charging operations depend on the configuration of the routes. In other words, the *tasks* that must be synchronized are not known a priori. The second difference is that in contrast to their problem, in our E-VRP-NL-C, the task synchronization has a direct impact on the solution cost. Dealing with these two features adds a thick layer of complexity to the models and solution methods.

The contribution of this paper is twofold. First, we propose two mixed integer linear programming (MILP) formulations for the E-VRP-NL-C, showing how capacity constraints can be integrated into standard E-VRP models. Second, we introduce a framework to solve E-VRPs with CS capacity constraints in general, and the E-VRP-NL-C in particular. It is made up of two interacting components: a *route generator* and a *solution assembler*. The first component builds a set of high-quality solutions, while relaxing the CS capacity constraints. The routes making up these solutions are stored in a pool which is sent to the assembler every few iterations. The latter combines routes from the pool in an attempt to build a solution meeting the CS capacity constraints. If such a solution does not exist or cannot be found within a given computing time, the assembler sends a signal to the generator to modify the search space in order to favor feasibility. For the particular case of the E-VRP-NL-C, we have developed a route generator based on a new, powerful, and efficient iterated local search (ILS) metaheuristic for the E-VRP-NL, and a solution assembler based on branch-and-cut. We present four assembly strategies allowing for different tradeoffs between efficiency and effectiveness. We have carried out extensive computational experiments on a large set of instances with different characteristics adapted from available benchmarks. The results demonstrate that our approach can handle the CS capacity constraints effectively. In addition, we report new BKS for 80 out of the 120 instances of a well-established benchmark set for the closely related E-VRP-NL.

The remainder of this paper is organized as follows. Section 2 formally introduces the E-VRP-NL-C. Section 3 describes MILP formulations of the problem. Section 4 presents the proposed solution method. Section 5 shows the computational results. Finally, Section 6 concludes and outlines research perspectives.

## 2 Problem description

We define the E-VRP-NL-C as follows. Let  $I$  be the set of customers that need to be served and let  $F$  be the set of charging stations (CSs) at which recharging can take place. Each customer  $i \in I$  has a service time  $g_i$ . The customers are served using a homogeneous fleet of EVs. Each EV has a battery of capacity  $Q$  (expressed in kWh). At the beginning of the planning horizon, the EVs are located in a single depot, from which they leave fully charged. The depot is continuously open for  $T_{max}$  hours. Traveling from one location  $i$  (the depot, a customer, or a CS) to another location  $j$  incurs a driving time  $t_{ij} \geq 0$  and an energy consumption  $e_{ij} \geq 0$ . Driving times and energy consumption both satisfy the triangular inequality. Due to their limited battery capacity, EVs may need to stop en route at CSs. Charging operations can occur at any CS, they are non-preemptive, and EVs can be partially recharged. The CS  $j \in F$  has a concave piecewise linear charging function  $\phi_j(\Delta)$  that maps for an empty battery the time  $\Delta$  spent charging at  $j$  and the SoC of the vehicle upon departing from  $j$ . If  $q$  is the SoC of the EV upon arrival at  $j$  and  $\Delta$  is the charging time, then the SoC of the EV upon departure from  $j$  is given by  $\phi_j(\Delta + \phi_j^{-1}(q))$  (Figure 1). We denote by  $B_j = \{0, \dots, b_j\}$  the ordered set of breakpoints of the charging function at  $j$ , sorted in non-decreasing order. We also introduce  $a_{jk}$  and  $q_{jk}$  to denote the charging time and the SoC for breakpoint  $k \in B_j$  of CS  $j$ .

Each CS  $j \in F$  also has a *capacity*, given by the number of available chargers  $C_j$ . Due to the limited capacity of CSs, vehicles may incur waiting times while they queue for a charger. We therefore note that optimal solutions are not

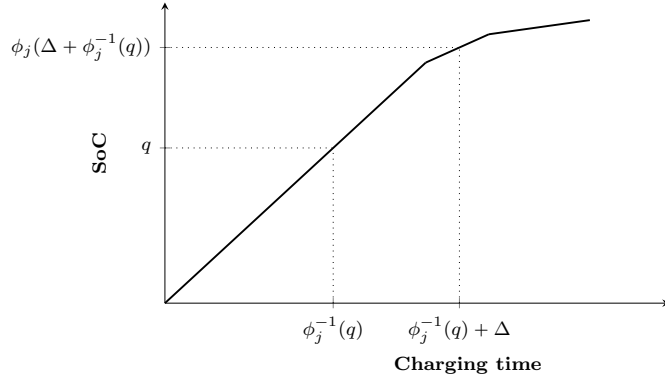


Figure 1: Piecewise linear charging function of a CS.

necessarily *left-shifted schedules*, as is the case for the E-VRP-NL and other E-VRP variants assuming uncapacitated CSs.

Feasible solutions to the E-VRP-NL-C must satisfy the following conditions: 1) each customer is visited exactly once by a single vehicle, 2) each route starts at the depot not before time 0 and ends at the depot not later than time  $T_{max}$ , 3) each route is energy-feasible, i.e., the SoC of an EV when it enters and leaves from any location lies between 0 and  $Q$ , and 4) no more than  $C_j$  EVs simultaneously charge at each CS  $j \in F$ . The objective of the E-VRP-NL-C is to minimize the total time needed to serve all customers, including driving, service, charging, and waiting times. Note that to avoid congestion at CSs, the starting times of the routes can be delayed at no cost. Moreover, an E-VRP-NL-C instance may have no feasible solution.

### 3 Two mixed-integer linear programming formulations

In this section we extend the formulations of Froger et al. (2019) for the closely related E-VRP-NL in order to deal with CS capacity constraints. These formulations belong to two families: CS replication-based formulations and path-based formulations. The former are akin to the MILP formulations which are typically used in the E-VRP literature. The latter correspond to a more intricate modeling strategy similar to the one proposed by Bartolini and Andelmin (2017) for the G-VRP. According to the results obtained by Froger et al. (2019) on small-size instances, path-based formulations outperform CS replication-based formulations. Nonetheless, we decided to explore both types of formulations to provide insights into how CS capacity constraints can be modeled under both families of formulations.

#### 3.1 CS replication-based formulation

The E-VRP-NL-C can be defined on a digraph  $G = (V, A)$ , where  $V = \{0\} \cup I \cup F'$  is the set of nodes and  $A$  is the set of arcs connecting the nodes of  $V$ . Node 0 represents the depot. The set  $F'$  is equal to  $\cup_{i \in F} F'_i$  where for each CS  $i \in F$   $F'_i$  denotes the set containing  $\beta_i$  copies of  $i$  (i.e.,  $|F'_i| = \beta_i$ ). The value of  $\beta_i$  corresponds to an upper bound on the number of visits to CS  $i$ . Note that we must have  $\beta_i \gg C_i$ ; otherwise, the capacity constraints are conservatively respected (the maximum number of charging operations at a CS being smaller than its capacity). We denote by  $F'_i \subseteq F'$  the set containing the  $\beta_i$  copies of CS  $i$  (i.e.,  $|F'_i| = \beta_i$  and  $F' = \cup_{i \in F} F'_i$ ). We assume that  $F'_i$  is an ordered set whose elements are numbered from 1 to  $\beta_i$ . In the remainder of this paper, depending on the context, we refer to an element of  $F'$  or  $F'_i$  as a CS copy or as a potential charging operation. We use the preprocessing procedure presented in Froger et al. (2019) to reduce the number of arcs in  $A$ . This technique primarily removes arcs that cannot feasibly be used due to the capacity of the battery.

According to the experiments carried out by Froger et al. (2019), the best CS replication-based formulation for the E-VRP-NL uses an arc-based tracking of the time and of the SoC. We therefore decided to use arc-based tracking in our first formulation. The binary variable  $x_{ij}$  is equal to 1 if and only if an EV travels on arc  $(i, j) \in A$ . The continuous variables  $\tau_{ij}$  and  $y_{ij}$  represent the time and SoC of an EV when it departs from vertex  $i \in V$  to travel on arc  $(i, j)$ . If no vehicle travels on this arc, then both variables are equal to 0. The continuous variables  $\underline{q}_j$  and  $\bar{q}_j$  specify the SoC of an

EV when it enters and leaves the CS copy  $j \in F'$ . The variables  $\underline{a}_j$  and  $\bar{a}_j$  are the scaled arrival and departure times, according to the charging function of CS copy  $j$ . The continuous variable  $\Delta_j$  represents the duration of the charging operation performed at  $j$ . For  $k \in B_j \setminus \{0\}$ , the binary variables  $\underline{w}_{jk}$  and  $\bar{w}_{jk}$  are equal to 1 if and only if the SoC lies between  $q_{j,k-1}$  and  $q_{jk}$  when the EV enters and leaves CS copy  $j$ , respectively. Finally, the continuous variables  $\underline{\lambda}_{jk}$  and  $\bar{\lambda}_{jk}$  are the coefficients associated with the breakpoints  $(a_{jk}, q_{jk})$  of the piecewise linear charging function  $\phi_j$  when the EV enters and leaves CS copy  $j$ , respectively. Without loss of generality, we restrict waiting times to occur only before charging operations. We introduce a continuous variable  $\nabla_j$  representing the waiting time of an EV before the start of its charging operation at  $j \in F'$ .

To model the CS capacity constraints, we rely on concepts borrowed from the Resource Constrained Scheduling Problem (RCPSP) literature. More precisely, we propose a *flow-based* formulation inspired from Artigues et al. (2003). There are, however, differences between our CS scheduling problem and the RCPSP. In our problem: *i*) the duration of each task (i.e., the charging operation), and *ii*) the number of tasks executed by each resource (i.e., a CS) are decision variables, whereas these are given parameters in the RCPSP. To the best of our knowledge this particular case has not yet been considered. In our flow-based formulation (hereafter referred to as FB), we consider  $C_i$  resources (represented by the chargers) for each CS  $i \in F$ . Each resource can execute at most one operation at any given time. Let  $i \in F$  be a CS and  $0_i$  and  $\beta_i + 1$  be two dummy operations acting as the source and the sink of the flow. We denote by  $\widetilde{F}'_i$  the ordered set  $\{0_i\} \cup F'_i \cup \{\beta_i + 1\}$  of potential charging operations (CS copies are considered here as potential charging operations). A visit to a CS copy requires the scheduling of the corresponding charging operation on one of the chargers of its associated CS. To break symmetries created by the introduction of copies of a CS  $i$ , without loss of generality, if  $\gamma_i \leq \beta_i$  visits to a CS are needed, we enforce the visit of the CS copies numbered from 1 to  $\gamma_i$ . Moreover, among the CS copies of  $i$  that are visited, we force them to be visited in the reverse order in which they appear in  $F'_i$  (i.e, a charging operation  $j \in F'_i$  must start after a charging operation  $l \in F'_i$  if  $l > j$ ). The reverse order is used since the departure time and charging duration variables are equal to zero when a CS copy is not visited. The sequential binary variable  $u_{jl}$  is equal to 1 only if (potential) operation  $l$  starts later than the completion of (potential) operation  $j$  with  $j, l \in F'_i$  and  $j > l$ . Let now  $j, l \in \widetilde{F}'_i$  be two operations (potentially dummy) such that  $j > l$ . The continuous flow variable  $f_{jl}$  denotes the number of chargers that are transferred from operation  $j$  to operation  $l$ . This number is naturally equal to one if  $j$  and  $l$  are not both dummy. Specifically, if both  $j$  and  $l$  are not dummy,  $f_{jl}$  is equal to one if and only if the operations  $j$  and  $l$  are scheduled on the same charger,  $l$  is scheduled after  $j$ , and no other operation uses the charger between the completion of  $j$  and the beginning of  $l$ . For notational convenience, we define  $\tilde{C}_j := 1$  for all  $j \in F'_i$  and  $\tilde{C}_{0_i} := \tilde{C}_{\beta_i+1} := C_i$ .

To illustrate our flow-based formulation, consider an example in which a CS  $i$  has two chargers (i.e.,  $C_i = 2$ ) and set  $F'_i$  contains five copies of this CS ( $\beta_i = 5$ ). Figure 2 depicts the structure of the flow network. Assume that the model decides to schedule four charging operations at  $i$ . Figure 3 illustrates the structure of the flow network derived from the schedule (by eliminating the flow variables associated to sequential variables equal to zero) and describes a feasible flow.

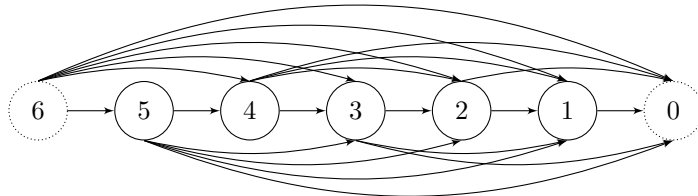
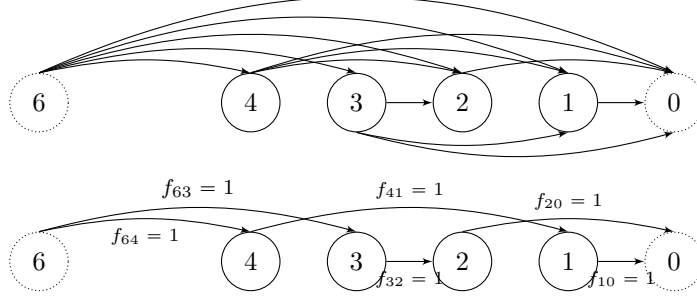
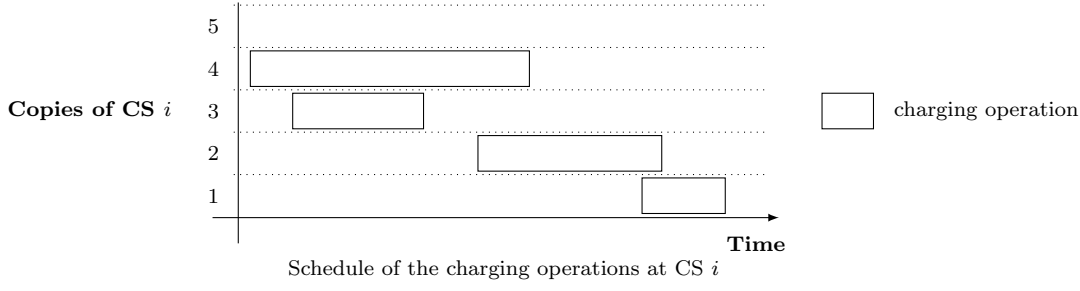


Figure 2: Structure of the flow network for CS  $i$ .



Structure of the flow network for CS  $i$  derived from the schedule (top) and a feasible flow (bottom)

Figure 3: Illustration of the flow-based formulation modeling approach for a given schedule at CS  $i$

Building on this concept, the CS replication-based formulation for the E-VRP-NL, denoted as  $[F^{CS_{rep}}]$ , is as follows:

$$[F^{CS_{rep}}] \text{ minimize } \sum_{(i,j) \in A} t_{ij} x_{ij} + \sum_{i \in F'} (\Delta_i + \nabla_i) \quad (1)$$

$$\text{subject to } \sum_{(i,l) \in A} x_{il} = 1 \quad i \in I \quad (2)$$

$$\sum_{(j,l) \in A} x_{jl} \leq 1 \quad j \in F' \quad (3)$$

$$\sum_{(l,i) \in A} x_{li} - \sum_{(i,l) \in A} x_{il} = 0 \quad i \in V \quad (4)$$

$$y_{0l} = Qx_{0l} \quad (0,l) \in A \quad (5)$$

$$\sum_{(l,i) \in A} y_{li} - \sum_{(l,i) \in A} e_{lj} x_{li} = \sum_{(i,l) \in A} y_{il} \quad i \in I \quad (6)$$

$$\sum_{(l,j) \in A} y_{lj} - \sum_{(l,j) \in A} e_{lj} x_{lj} = \underline{q}_j \quad j \in F' \quad (7)$$

$$\sum_{(j,l) \in A} y_{jl} = \bar{q}_j \quad j \in F' \quad (8)$$

$$y_{ij} \leq \left( Q - \min_{l \in F' \cup \{0\}} \{e_{li}\} \right) x_{ij} \quad (i,j) \in A \quad (9)$$

$$y_{ij} \geq \left( e_{ij} + \min_{l \in F' \cup \{0\}} \{e_{jl}\} \right) x_{ij} \quad (i,j) \in A \quad (10)$$

$$\underline{q}_j \leq \bar{q}_j \quad j \in F' \quad (11)$$

$$\underline{q}_j = \sum_{k \in B_j} \lambda_{jk} \underline{q}_{jk} \quad j \in F' \quad (12)$$

$$\underline{a}_j = \sum_{k \in B_j} \lambda_{jk} \underline{a}_{jk} \quad j \in F' \quad (13)$$

$$\sum_{k \in B_j} \lambda_{jk} = \sum_{k \in B_j \setminus \{0\}} \underline{w}_{jk} \quad j \in F' \quad (14)$$

$$\sum_{k \in B_j \setminus \{0\}} \underline{w}_{jk} = \sum_{(j,l) \in A} x_{jl} \quad j \in F' \quad (15)$$

$$\lambda_{j0} \leq \underline{w}_{j1} \quad j \in F' \quad (16)$$

$$\underline{\lambda}_{jk} \leq \underline{w}_{jk} + \underline{w}_{j,k+1} \quad j \in F', k \in B_j \setminus \{0, b_j\} \quad (17)$$

$$\underline{\lambda}_{jb_j} \leq \underline{w}_{jb_j} \quad j \in F' \quad (18)$$

$$\bar{q}_j = \sum_{k \in B_j} \bar{\lambda}_{jk} q_{jk} \quad j \in F' \quad (19)$$

$$\bar{a}_j = \sum_{k \in B_j} \bar{\lambda}_{jk} a_{jk} \quad j \in F' \quad (20)$$

$$\sum_{k \in B_j} \bar{\lambda}_{jk} = \sum_{k \in B_j \setminus \{0\}} \bar{w}_{jk} \quad j \in F' \quad (21)$$

$$\sum_{k \in B_j \setminus \{0\}} \bar{w}_{jk} = \sum_{(j,l) \in A} x_{jl} \quad j \in F' \quad (22)$$

$$\bar{\lambda}_{j0} \leq \bar{w}_{j1} \quad j \in F' \quad (23)$$

$$\bar{\lambda}_{jk} \leq \bar{w}_{jk} + \bar{w}_{j,k+1} \quad j \in F', k \in B_j \setminus \{0, b_j\} \quad (24)$$

$$\bar{\lambda}_{jb_j} \leq \bar{w}_{jb_j} \quad j \in F' \quad (25)$$

$$\Delta_j = \bar{a}_j - \underline{a}_j \quad j \in F' \quad (26)$$

$$\sum_{(l,i) \in A} (\tau_{li} + (t_{li} + g_i) x_{li}) = \sum_{(i,l) \in A} \tau_{il} \quad i \in I \quad (27)$$

$$\sum_{(l,j) \in A} (\tau_{lj} + t_{lj} x_{lj}) + \Delta_j + \nabla_j = \sum_{(j,l) \in A} \tau_{jl} \quad j \in F' \quad (28)$$

$$\tau_{li} \leq (T_{max} - t_{li} - g_i - t_{i0}) x_{li} \quad (l,i) \in A, i \in I \quad (29)$$

$$\tau_{lj} \leq (T_{max} - t_{lj} - \Delta_j^{min} - t_{j0}) x_{lj} \quad (l,j) \in A, j \in F' \quad (30)$$

$$\sum_{(h,j) \in A} x_{hj} \leq \sum_{(h,l) \in A} x_{hl}, \quad i \in F, j, l \in F'_i, j < l \quad (31)$$

$$\sum_{(j,h) \in A} \tau_{jh} - \Delta_j \geq \sum_{(l,h) \in A} \tau_{lh} - \Delta_l \quad i \in F, j, l \in F'_i : j < l \quad (32)$$

$$\sum_{l \in \tilde{F}'_i, l > j} f_{lj} = \sum_{(l,j) \in A} x_{lj} \quad i \in F, j \in F'_i \quad (33)$$

$$\sum_{l \in \tilde{F}'_i, l > j} f_{lj} - \sum_{l \in \tilde{F}'_i, l < j} f_{jl} = 0 \quad i \in F, j \in F'_i \quad (34)$$

$$C_i - \sum_{l \in \tilde{F}'_i, l > 0_i} f_{l,0_i} = 0 \quad i \in F \quad (35)$$

$$\sum_{l \in \tilde{F}'_i, l < \beta_i + 1} f_{\beta_i + 1, l} - C_i = 0 \quad i \in F \quad (36)$$

$$u_{jh} \geq u_{jl} + u_{lh} - 1 \quad i \in F, j, l, h \in F'_i : j > l > h \quad (37)$$

$$\sum_{(l,h) \in A} \tau_{lh} - \Delta_l - \sum_{(j,h) \in A} \tau_{jh} \geq (T_{max} - t_{i0}) (u_{jl} - 1) \quad i \in F, j, l \in F'_i, j > l \quad (38)$$

$$f_{jl} \leq \min(\tilde{C}_j, \tilde{C}_l) u_{jl} \quad i \in F, (j,l) \in F'_i, j > l \quad (39)$$

$$x_{ij} \in \{0, 1\}, \tau_{ij} \geq 0, y_{ij} \geq 0 \quad (i,j) \in A \quad (40)$$

$$\underline{w}_{jk} \in \{0, 1\}, \bar{w}_{jk} \in \{0, 1\} \quad j \in F', k \in B_j \setminus \{0\} \quad (41)$$

$$\underline{\lambda}_{jk} \geq 0, \bar{\lambda}_{jk} \geq 0 \quad j \in F', k \in B_j \quad (42)$$

$$\underline{a}_j \geq 0, \bar{a}_j \geq 0, \underline{q}_j \geq 0, \bar{q}_j \geq 0, \Delta_j \geq 0, \nabla_j \geq 0 \quad j \in F' \quad (43)$$

$$u_{jl} \in \{0, 1\} \quad i \in F, j, l \in F'_i, j > l \quad (44)$$

$$f_{jl} \geq 0 \quad i \in F, j, l \in \tilde{F}'_i, j > l. \quad (45)$$

The objective function (1) includes the driving, charging, and waiting times. Constraints (2) ensure that each customer is visited exactly once, while constraints (3) ensure that each CS copy is visited at most once. Constraints (4) impose the flow conservation. Constraints (5) state that the EVs leave the depot fully charged. Constraints (6) track the SoC at each customer. Constraints (7) track the SoC of the EV when it arrives at a CS copy. Constraints (8) track the SoC of the EV when it leaves a CS copy. Constraints (9) couple the  $y_{ij}$  and  $x_{ij}$  variables. Constraints (10) state that if an



EV traverses the arc  $(i, j)$  its SoC when leaving  $i$  must be enough to traverse the arc and then to reach the closest CS or the depot. Constraints (11) couple the SoC of an EV upon arrival at a CS with its SoC at departure. Constraints (12)–(18) define the SoC and the corresponding charging time with respect to  $\phi_j$  of an EV upon arrival at CS copy  $j$ . Specifically, we express  $(\underline{a}_j, \underline{q}_j)$  as a convex combination of the breakpoints  $(a_{jk}, q_{jk})$  using the  $\underline{\lambda}_{jk}$  variables, where  $k \in B_j$ . Similarly, constraints (19)–(25) define the SoC and the corresponding charging time with respect to  $\phi_j$  of an EV upon departure from CS copy  $j$ . Specifically, we express  $(\bar{a}_j, \bar{q}_j)$  as a convex combination of the breakpoints  $(a_{jk}, q_{jk})$  using the  $\bar{\lambda}_{jk}$  variables, where  $k \in B_j$ . Constraints (26) define the time spent at a CS copy. Constraints (27) track the departure time at each customer. Constraints (28) track the departure time at CS copies. Constraints (29) and (30) couple the  $\tau_{ij}$  and  $x_{ij}$  variables. Specifically, if an EV traverses an arc  $(i, j)$ , then its departure time must guarantee that the EV returns to the depot before the closing time. The parameter  $\Delta_j^{min}$  is a lower bound on the time spent charging to recover the energy consumed to make the detour to  $j \in F'$ . It is equal to the time needed to charge the energy quantity  $\min_{l, l' \in V \setminus \{j\}; (l, j) \in A \wedge (j, l') \in A} \{e_{lj} + e_{jl'} - e_{ll'}\}$  if the EV has an empty battery. Constraints (31) and (32) break the symmetries created by the introduction of the CS copies. Constraints (33) state that a resource has to be allocated to a charging operation in  $F'_i$  if an EV reaches the corresponding CS. Constraints (34)–(36) ensure the flow conservation for the chargers. Constraints (37) express the transitivity of the precedence relations. Constraints (38) are the disjunctive constraints coupling the start time of charging operations  $j$  and  $l$  to  $u_{jl}$ . For a pair  $(j, l)$ , the constraint is active when  $u_{jl} = 1$  and, in this case, it enforces the precedence relation between the charging operations  $j$  and  $l$  (i.e.,  $l$  cannot start before the completion of  $j$ ). Constraints (39) couple the flow variables for the chargers to the charging operation sequence variables. Finally, constraints (40)–(45) define the domains of the decision variables.

### 3.2 Path-based formulation

One drawback of the previous formulation is the need to replicate the CSs. To ensure that no optimal solution is cut off, the number of needed copies is very large (Froger et al. (2019) provide an example where this value is equal to  $4|I|$ ), which yields intractable MILPs. To overcome this difficulty, Froger et al. (2019) proposed an alternative model of for the E-VRP-NL based on the concept of CS paths between each couple of nodes (either customers or the depot). Given two nodes (customer or depot)  $i$  and  $j$ , we call a CS path (CSP) a simple path starting from  $i$  visiting a sequence of CSs or none and ending at  $j$ .

The concept of CSPs leads to a redefinition of the problem on a directed multigraph  $\tilde{G} = (\tilde{V}, \tilde{A})$ , where  $\tilde{V} = \{0\} \cup I$ , and  $\tilde{A}$  is the set of arcs connecting the nodes of  $\tilde{V}$ . More specifically, an arc in  $\tilde{A}$  represents a CSP  $p$ , starting in  $\mathbf{org}(p) \in \tilde{V}$  and ending in  $\mathbf{dest}(p) \in \tilde{V}$ . We denote  $n_p$  the number of CSs in  $p$  and  $L_p = \{0, \dots, n_p - 1\}$  as the ordered set of CS positions in  $p$ . We note that  $L_p = \emptyset$  indicates that  $p$  does not visit any CS between  $\mathbf{org}(p)$  and  $\mathbf{dest}(p)$ . When  $L_p \neq \emptyset$ , we denote  $\mu_p(l)$  the CS at position  $l \in L_p$ . We define  $e^p$  and  $t^p$  as the energy consumption and the driving time associated with CSP  $p \in P$ . Given two nodes  $i, j \in \tilde{V}$ , we define  $P_{ij}$  as the set of CSPs connecting  $i$  to  $j$ , and we define  $P = \bigcup_{i, j \in \tilde{V}, i \neq j} P_{ij}$  as the set of all CSPs.

Our path-based formulation of the E-VRP-NL-C involves the following decisions variables. The binary variable  $x_p$  is 1 if and only if an EV travels CSP  $p \in P$ . The continuous variables  $\tau_p$  and  $y_p$  track the time and SoC of an EV when it departs from node  $\mathbf{org}(p)$  to  $\mathbf{dest}(p)$  using CSP  $p$ . We associate with each CS in a CSP of set  $P$  a unique potential charging operation (a CSP can at most be traveled by a single EV). The continuous variables  $\underline{q}_{pl}$  and  $\bar{q}_{pl}$  specify the SoC of an EV when it enters and leaves  $\mu_p(l)$  (i.e., the CS at position  $l \in L_p$ ). The continuous variable  $\Delta_{pl}$  represents the duration of the charging operation performed at  $\mu_p(l)$ . The variables  $\underline{a}_{pl}$  and  $\bar{a}_{pl}$  are the scaled arrival and departure times, according to the charging function of CS  $\mu_p(l)$ . For  $k \in B_i \setminus \{0\}$ , the binary variables  $\underline{w}_{plk}$  and  $\bar{w}_{plk}$  are equal to 1 if and only if the SoC is between  $q_{\mu_p(l), k-1}$  and  $q_{\mu_p(l), k}$  when the EV enters and leaves CS  $\mu_p(l)$ , respectively. The continuous variables  $\underline{\lambda}_{plk}$  and  $\bar{\lambda}_{plk}$  represent the coefficients associated with the breakpoints  $(a_{\mu_p(l), k}, q_{\mu_p(l), k})$  of the piecewise linear charging function, when the EV enters and leaves CS  $\mu_p(l)$ , respectively. Let  $e^p$  and  $t^p$  be the energy consumption and the driving time associated with CSP  $p \in P$ . The continuous variable  $\nabla_{pl}$  is the waiting time at the CS at position  $l$  in CSP  $p$  before charging. We also need to track the starting and completion time of every charging operation. The continuous variables  $\underline{s}_{pl}$  and  $\bar{s}_{pl}$  represent the starting and completion time of the charging operation performed at  $\mu_p(l)$ .

The modeling of the capacity constraints follows the same idea as the one described in §3.1. For convenience, we introduce the set  $O_i$  of potential charging operations at CS  $i \in F$ . Every operation  $o \in O_i$  represents the visit of a

specific CS of a CSP of  $P$ . We denote by  $\bar{p}(o) \in P$  the corresponding CSP and by  $l(o) \in L_{\bar{p}(o)}$  the position of the visited CS. Specifically, an operation  $o$  is associated with CS  $\mu_{\bar{p}(o)}(l(o))$ . Binary variable  $u_{oo'}$  is equal to 1 if and only if operation  $o'$  starts after the completion of operation  $o \neq o'$ . For every CS  $i \in F$ , we introduce two dummy operations  $\varepsilon_i^+$  and  $\varepsilon_i^-$ , which act as the source and the sink of the flow. Let  $(o, o') \in (O_i \cup \{\varepsilon_i^+\}) \times (O_i \cup \{\varepsilon_i^-\})$  be a couple of charging operations; then the continuous flow variable  $f_{oo'}$  denotes the number of chargers that are transferred from operation  $o$  to operation  $o'$ . When both these operations are not dummy,  $f_{oo'}$  is equal to one if and only if these operations are scheduled on the same charger,  $o'$  is scheduled after  $o$ , and no other operation is scheduled on the charger between the completion of  $o$  and the beginning of  $o'$ . For notational convenience, we define  $\tilde{C}_o := 1$  for all  $o \in O_i$  and  $\tilde{C}_{\varepsilon_i^+} := \tilde{C}_{\varepsilon_i^-} := C_i$ . The second model for the E-VRP-NL-C, denoted as  $[P^{path}]$ , is as follows:

$$[P^{path}] \quad \text{minimize} \quad \sum_{p \in P} \left( t^p x_p + \sum_{l \in L_p} (\Delta_{pl} + \nabla_{pl}) \right) \quad (46)$$

$$\text{subject to} \quad \sum_{j \in \tilde{V}, i \neq j} \sum_{p \in P_{ij}} x_p = 1 \quad i \in I \quad (47)$$

$$\sum_{j \in \tilde{V}, i \neq j} \sum_{p \in P_{ji}} x_p - \sum_{j \in \tilde{V}, i \neq j} \sum_{p \in P_{ij}} x_p = 0 \quad i \in \tilde{V} \quad (48)$$

$$\sum_{l \in \tilde{V}, l \neq j} \sum_{p \in P_{lj}} \left( y_p - e^p x_p + \sum_{l \in L_p} (\bar{q}_{pl} - \underline{q}_{pl}) \right) = \sum_{l \in \tilde{V}, l \neq j} \sum_{p \in P_{jl}} y_p \quad j \in I \quad (49)$$

$$y_p - e_{\text{org}(p), \mu_p(0)} x_p = \underline{q}_{p0} \quad p \in P \quad (50)$$

$$\bar{q}_{p, l-1} - e_{\mu_p(l-1), \mu_p(l)} x_p = \underline{q}_{pl} \quad p \in P, l \in L_p \setminus \{0\} \quad (51)$$

$$y_p - e^p x_p - \sum_{l \in L_p} (o_{pl} - q_{pl}) \geq 0 \quad i \in I, p \in P_{i0} \quad (52)$$

$$y_p = Q x_p \quad i \in \tilde{V} \setminus \{0\}, p \in P_{0i} \quad (53)$$

$$y_p \leq Q x_p \quad p \in P \quad (54)$$

$$\underline{q}_{pl} = \sum_{k \in B_{\mu_p(l)}} \underline{\lambda}_{plk} q_{\mu_p(l)k} \quad p \in P, l \in L_p \quad (55)$$

$$\underline{a}_{pl} = \sum_{k \in B_{\mu_p(l)}} \underline{\lambda}_{plk} a_{\mu_p(l)k} \quad p \in P, l \in L_p \quad (56)$$

$$\sum_{k \in B_{\mu_p(l)}} \underline{\lambda}_{plk} = \sum_{k \in B_{\mu_p(l)} \setminus \{0\}} \underline{w}_{plk} \quad p \in P, l \in L_p \quad (57)$$

$$\sum_{k \in B_{\mu_p(l)} \setminus \{0\}} \underline{w}_{plk} = x_p \quad p \in P, l \in L_p \quad (58)$$

$$\underline{\lambda}_{pl0} \leq \underline{w}_{pl1} \quad p \in P, l \in L_p \quad (59)$$

$$\underline{\lambda}_{plk} \leq \underline{w}_{plk} + \underline{w}_{pl, k+1} \quad p \in P, l \in L_p, k \in B_{\mu_p(l)} \setminus \{0, b_{\mu_p(l)}\} \quad (60)$$

$$\underline{\lambda}_{plb_{\mu_p(l)}} \leq \underline{w}_{plb_{\mu_p(l)}} \quad p \in P, l \in L_p \quad (61)$$

$$\bar{q}_{pl} = \sum_{k \in B_{\mu_p(l)}} \bar{\lambda}_{plk} q_{\mu_p(l)k} \quad p \in P, l \in L_p \quad (62)$$

$$\bar{a}_{pl} = \sum_{k \in B_{\mu_p(l)}} \bar{\lambda}_{plk} a_{\mu_p(l)k} \quad p \in P, l \in L_p \quad (63)$$

$$\sum_{k \in B_{\mu_p(l)}} \bar{\lambda}_{plk} = \sum_{k \in B_{\mu_p(l)} \setminus \{0\}} \bar{w}_{plk} \quad p \in P, l \in L_p \quad (64)$$

$$\sum_{k \in B_{\mu_p(l)} \setminus \{0\}} \bar{w}_{plk} = x_p \quad p \in P, l \in L_p \quad (65)$$

$$\bar{\lambda}_{i0} \leq \bar{w}_{pl1} \quad p \in P, l \in L_p \quad (66)$$

$$\bar{\lambda}_{plk} \leq \bar{w}_{plk} + \bar{w}_{pl, k+1} \quad p \in P, l \in L_p, k \in B_{\mu_p(l)} \setminus \{0, b_{\mu_p(l)}\} \quad (67)$$

$$\bar{\lambda}_{plb_{\mu_p(l)}} \leq \bar{w}_{plb_{\mu_p(l)}} \quad p \in P, l \in L_p \quad (68)$$

$$\Delta_{pl} = \bar{a}_{pl} - \underline{a}_{pl} \quad p \in P, l \in L_p \quad (69)$$

$$\sum_{i \in V, i \neq j} \sum_{p \in P_{ij}} \left( \tau_p + t^p x_p + \sum_{l \in L_p} (\Delta_{pl} + \nabla_{pl}) \right) + g_j = \sum_{l \in \bar{V}, l \neq j} \sum_{p \in P_{jl}} \tau_p \quad j \in I \quad (70)$$

$$\tau_p + \sum_{l \in L_p} (\Delta_{pl} + \nabla_{pl}) \leq (T_{max} - t^p - p_{dest(p)} - t_{dest(p),0}) x_p \quad p \in P \quad (71)$$

$$\tau_p + t_{org(p),\mu_p(0)} x_p + \nabla_{p0} = \underline{s}_{p0} \quad p \in P \quad (72)$$

$$\bar{s}_{p,l-1} + t_{\mu_p(l-1),\mu_p(l)} x_p + \nabla_{pl} = \underline{s}_{pl} \quad p \in P, l \in L_p, l \neq 0 \quad (73)$$

$$\Delta_{pl} = \bar{s}_{pl} - \underline{s}_{pl} \quad p \in P, l \in L_p \quad (74)$$

$$\sum_{o' \in O_i \cup \{\varepsilon_i^+\}} f_{o'o} = x_{p(o)} \quad i \in F, o \in O_i \quad (75)$$

$$\sum_{o' \in O_i \cup \{\varepsilon_i^+\}} f_{o'o} - \sum_{o' \in O_i \cup \{\varepsilon_i^-\}} f_{oo'} = 0 \quad i \in F, o \in O_i \quad (76)$$

$$C_i - \sum_{o \in O_i \cup \{\varepsilon_i^-\}} f_{\varepsilon_i^+,o} = 0 \quad i \in F \quad (77)$$

$$\sum_{o \in O_i \cup \{\varepsilon_i^+\}} f_{o,\varepsilon_i^-} - C_i = 0 \quad i \in F \quad (78)$$

$$\underline{s}_{p(o),l(o)} - \bar{s}_{p(o'),l(o')} \geq T_{max} (u_{o'o} - 1) \quad i \in F, o, o' \in O_i \quad (79)$$

$$f_{oo'} \leq \min(\tilde{C}_o, \tilde{C}_{o'}) u_{oo'} \quad i \in F, (o, o') \in (O_i \cup \{\varepsilon_i^+\}) \times (O_i \cup \{\varepsilon_i^-\}) \quad (80)$$

$$x_p \in \{0, 1\} \quad p \in P \quad (81)$$

$$\tau_p \geq 0, y_p \geq 0 \quad p \in P \quad (82)$$

$$\underline{q}_{pl}, \bar{q}_{pl}, \underline{a}_{pl}, \bar{a}_{pl}, \underline{s}_{pl}, \bar{s}_{pl}, \Delta_{pl}, \nabla_{pl} \geq 0 \quad p \in P, l \in L_p \quad (83)$$

$$\underline{\lambda}_{plk} \geq 0, \bar{\lambda}_{plk} \geq 0 \quad p \in P, l \in L_p, k \in B_{\mu_p(l)} \setminus \{0\} \quad (84)$$

$$\underline{w}_{plk} \in \{0, 1\}, \bar{w}_{plk} \in \{0, 1\} \quad p \in P, l \in L_p, k \in B_{i(p,l)} \setminus \{0\} \quad (85)$$

$$u_{oo'} \in \{0, 1\} \quad i \in F, o, o' \in O_i \quad (86)$$

$$f_{oo'} \geq 0 \quad i \in F, (o, o') \in (O_i \cup \{\varepsilon_i^+\}) \times (O_i \cup \{\varepsilon_i^-\}). \quad (87)$$

The objective function (46) minimizes the total driving, charging, and waiting time. Constraints (47) ensure that each customer is visited exactly once. Constraints (48) impose flow conservation. Constraints (49) track the SoC of EVs at each customer location. Constraints (50) track the SoC at the arrival at the first CS of each CSP. Constraints (51) couple the SoC of an EV that leaves a CS to go to another CS. Constraints (52) ensure that if the EV travels between a vertex and the depot, it has sufficient energy to reach its destination. Constraints (53) state that every EV leaves the depot with a fully charged battery. Constraints (54) couple the SoC tracking variable with the arc travel variables. Constraints (55)–(69) model the relationship between the SoC of an EV upon arrival at a CS, the charging time, and its SoC upon departing from this CS. Constraints (70) track the departure time at each vertex. Constraints (71) couple the time tracking variable with the arc travel variables, and impose the route duration limit. Constraints (72) and (73) define the starting and completion times of every charging operation, as well as the potential waiting time before the start of the operation. Constraints (74) define the duration of the charging operations based on their starting and completion times. Constraints (75) state that a charger has to be allocated to every charging operation of each selected CSP. Constraints (76) ensure flow conservation. Constraints (77) and (78) compute the flow value at the beginning and at the end of the time horizon. Constraints (79) couple the sequencing variables with the starting time of the corresponding charging operations. Constraints (80) couple the flow variables with the sequence variables. Specifically, a charger can be sent from a charging operation  $o'$  to another charging operation  $o$  if  $o$  starts after the completion of  $o'$ . Finally, constraints (81)–(87) define the domains of the decision variables.

Without preprocessing, the number of CSPs explodes with the number of CSs and the number of customers. However, a large number of these arcs cannot be part of an optimal solution. Froger et al. (2019) presented a filtering procedure to reduce the number of CSPs in the path-based formulation based on the definition of a dominance rule between two CSPs having the same origin and destination. Due to the potential waiting times that can occur at CSs, we cannot apply the dominance rule described in the work of Froger et al. (2019) between CSPs with the same origin and destination if they both contain CSs. However, in our computational experiments we apply the dominance rule in the special case

between the unique CSP with no CS and every other CSP of  $P_{ij}$ , with  $i$  and  $j$  in  $\tilde{V}$ .

## 4 Solution method

The results obtained by Froger et al. (2019) suggest that the models introduced in §3.1 are only applicable to small-size instances (around 20 customers). In this section, we introduce a matheuristic to solve the E-VRP-NL-C. The method relies on two interacting components: a route generator and a solution assembler. The first component builds a pool of high-quality solutions by relaxing the CS capacity constraints, while the latter component recombines routes from the pool trying to build a solution satisfying the CS capacity constraints. Algorithm 1 outlines the general structure of the method. The algorithm starts by generating an initial solution without taking into account the capacity constraints (line 1). In our implementation, this step is carried out using a modified version of the classical Clarke and Wright heuristic. Next, the algorithm enters the main loop (lines 3–28). During  $n_{max}$  iterations the algorithm alternates between the *route generation* and *solution assembly* phases. In particular, the algorithm uses procedure `generateRoutes(·)` to retrieve a triplet  $(\Phi_O, \Phi_R, \Omega')$ , where  $\Phi_O, \Phi_R$  are sets of high-quality solutions for the *original* and *relaxed* problems (i.e., with and without CS capacity constraints), and  $\Omega'$  is a set of independent, feasible, and high-quality routes for the original problem. The latter will be referred to as the *short-term* pool. Next (lines 4–12) the algorithm adds the routes making up the solutions in  $\Phi_O$  and  $\Phi_R$  to a *long-term* pool  $\Omega$  while keeping track of the best solutions for both the original ( $s_O^*$ ) and the relaxed problem ( $s_R^*$ ). The intuition behind adding routes coming from potentially infeasible solutions to the original problem (i.e., solutions to the relaxed problem) to the long-term pool is to foster diversity. Indeed, these routes tend to be of high quality (in terms of the objective function) and might be recombined efficiently with others later.

The algorithm then enters the assembly phase (lines 13–14). In the first step of this phase, the algorithm merges the short- and long-term pools. Thus, seeking to combine the past and recent history of the search into a single set of routes that has a size that is manageable for the solution assembler. The algorithm then calls procedure `assemble(·)` to retrieve a tuple  $(s', S')$ , where  $s'$  is the best solution and  $S'$  is the set of all improving solutions (with respect to  $s_O^*$ ) for the original problem found during the call. If the assembler retrieves a solution, the algorithm adds its routes in  $S'$  to the long-term pool  $\Omega$  and updates the incumbent  $s_O^*$  (lines 15–18). If after the call to the assembler, the algorithm still has not found a feasible solution (i.e.,  $s_O^*$  is still equal to `null`), it implements two actions. First, it slightly modifies the search space to favor feasibility. In our implementation, we modify the solution space by artificially reducing the opening hours of the depot (line 21). The underlying idea, is that the reduction of the depot’s operating hours would lead to shorter routes with more slack to accommodate waiting times at charging stations during the assembly phase. Second, it tries to *repair* the best-known solution for the relaxed problem (i.e.,  $s_R^*$ ) by making it comply with the new solution space. In our implementation, if a route visiting  $n$  customers has a duration strictly greater than the new value  $T$  of the depot hours, the algorithm creates two new routes by adding a return to the depot after the  $\lfloor n/2 \rfloor^{th}$  customer and reoptimizes the charging decisions within these routes. To avoid feasibility issues, the new route maximum duration limit is not considered when a route visits only one customer. The same procedure is performed on the newly created routes as long as the routes contain more than one customer and their duration exceeds  $T$ . On the other hand, if after the call to the assembler a feasible solution is available (i.e.,  $s_O^*$  is different than `null`), the algorithm then reestablishes (if needed) the original solution space (i.e., resets the value of the depot opening hours to  $T_{max}$ ) and moves to a new iteration.

Both the route generation and solution assembly phases consist in solving complex combinatorial optimisation problems. In our framework, we develop an iterative local search algorithm as the route generator and a branch-and-cut procedure as the solution assembler. The remainder of this section describes these two components.

### 4.1 Route generator: an iterated local search algorithm for the E-VRP-NL

Local search (LS) techniques have been commonly used to solve E-VRPs. In this respect, researchers have devoted substantial effort in finding move evaluation strategies that offer the best trade-off between accuracy and efficiency. A first approach relies on LS operators that focus on customers or CSs, more or less indistinctively, while other approaches focus on one type of node (i.e., either customers or CSs) while not modifying the positioning of the others ones. Typically, given the resulting sequence of customers and CSs, the quantity to be charged at each CS is determined. The feasibility and the evaluation of a move is therefore a result of this process. While this has the advantage of

---

**Algorithm 1:** Solution method - general structure

---

```
/*  $f(s)$  denotes the value of the objective function for a solution  $s$  and we assume that  $f(\text{NULL}) = +\infty$  */
1  $s_R^0 \leftarrow \text{generateInitialSolution}()$ 
2  $n \leftarrow 0, T \leftarrow T_{max}, \Omega \leftarrow \emptyset, s_R^* \leftarrow s_R^0, s_R \leftarrow s_R^0, s_{FO}^* \leftarrow \text{NULL}, s_O \leftarrow \text{NULL}$ 
3 while  $n < n_{max}$  do
4    $(\Phi_O, \Phi_R, \Omega') \leftarrow \text{generateRoutes}(s_R, s_O, \delta, T)$  (see Algorithm 2)
5   for each  $s'_R \in \Phi_R$  do
6      $\text{addRoutesToPool}(\Omega, s'_R)$ 
7     if  $f(s'_R) < f(s_R^*)$  then  $s_R^* \leftarrow s'_R$ 
8   end
9   for each  $s'_O \in \Phi_O$  do
10     $\text{addRoutesToPool}(\Omega, s'_O)$ 
11    if  $f(s'_O) < f(s_O^*)$  then  $s_O^* \leftarrow s'_O$ 
12  end
13   $\Omega' \leftarrow \text{merge}(\Omega, \Omega')$ 
14   $(s', S') \leftarrow \text{assemble}(\Omega', s_O^*)$  (see Algorithm 4) /*  $s' = \text{NULL}$  if no improving solution is found */
15  if  $s' \neq \text{NULL}$  then
16     $s_O^* \leftarrow s'$ 
17    if  $f(s') < f(s_R^*)$  then  $s_R^* \leftarrow s'$ 
18    for each  $s \in S'$  do  $\text{addRoutesToPool}(\Omega, s)$ 
19  end
20  if  $s_O^* = \text{NULL}$  then
21     $T \leftarrow \max\{T_{min}, \alpha \cdot T\}$  /*  $T_{min}$  : minimum possible value for  $T$ ,  $\alpha < 1$  a tuning parameter */
22     $s_R \leftarrow \text{repair}(s_R^*), s_O \leftarrow \text{NULL}$ 
23  else
24    if  $T < T_{max}$  then  $T \leftarrow T_{max}$ 
25     $s_O \leftarrow s_O^*, s_R \leftarrow s_R^*$ 
26  end
27   $n \leftarrow n + 1$ 
28 end
29 return  $s_O^*$ 
```

---

requiring a low computational effort, it can have undesired effects. For example, a move may be deemed infeasible or non-improving, while the resulting sequence(s) of customer visits can lead to an excellent solution if the charging decisions are reoptimized, in terms of which CSs to visit, when to visit them and how much to charge the EV during these visits.

Most of the LS-based algorithms designed for E-VRPs use the aforementioned approaches (e.g., Felipe et al. (2014); Schneider et al. (2014); Goeke and Schneider (2015); Koç and Karaoglan (2016)). In some cases, more advanced strategies are also called periodically to improve the charging decisions inside the routes, by means of heuristic or exact procedures, or a combination of both (Hiermann et al. (2016); Montoya et al. (2017)). Another approach couples the charging decisions (in terms of which CSs to visit, when to visit them and how much to charge the EV during these visits) with the evaluation of the customer sequencing moves. Specifically, it relies on LS operators that solely function on customers, while the charging decisions are only made to check the feasibility of a move and to evaluate it. This coupling approach comes with a computational burden. To avoid the need to explicitly model CSs, Andelmin and Bartolini (2019) designed for the Green VRP a multi-start local search heuristic that works on a multigraph (where each arc represents a CSP). Nonetheless, their solution method is only halfway between the previously described approaches for two reasons. First, vehicles are fully charged when visiting a CS, and thus no optimisation of the recharging quantities is performed. Second, they do not necessarily guarantee that the charging decisions inside the routes are optimal. To our knowledge, only Hiermann et al. (2019) designed an LS-based algorithm based on the coupling approach. Specifically, they used a dynamic programming procedure to determine for each single route the CS visits and greedy policies to decide the propulsion mode for a plug-in hybrid electric vehicle. They call this procedure every time they evaluate a move.

The previous discussion shows that charging decisions should be handled with care in LS-based algorithms for E-VRPs. In the E-VRP-NL-C, given a set of routes without CSs, making optimal charging decisions consists in defining

for every route where, when and how much to charge in order to minimize the total time to visit the customers while meeting the CS capacity constraints. Therefore, the charging decisions of the routes need to be coupled, thus adding another complicating decisional layer. As a consequence, solving this charging problem whenever an LS-based algorithm evaluates a move is computationally demanding for the E-VRP-NL-C. However, for the vast majority of E-VRPs this computational burden can be reduced since charging decisions can be independently optimized for each route (defined here as a sequence of customer visits). This optimisation problem was explicitly handled by Montoya et al. (2017) and Froger et al. (2019) in the context of the E-VRP-NL. The authors referred to this problem as the fixed route vehicle charging problem (FRVCP).

We have developed a novel route generator, which is based on an ILS algorithm that solves the E-VRP-NL, while populating a pool of routes to be used in constructing solutions to the E-VRP-NL-C. In particular, we considered only customer nodes in the local search, and we used the labeling algorithm of Froger et al. (2019) to tackle the FRVCP, as their results suggest that it can be embedded within heuristics.

Introduced by Lourenço et al. (2003), ILS is a metaheuristic that iteratively applies an LS phase to produce local optima, and perturbation mechanisms to escape from them. It is initialized with a solution generally provided by a constructive heuristic. In our implementation, we combine ILS with a variable neighborhood descent (VND) search strategy for the LS phase. Algorithm 2 outlines the general structure of our method. First, the current best solution  $s_R^*$  is set equal to the initial solution  $s_R^0$  provided as input. The algorithm then enters an iterative process. Except during the first iteration, it perturbs the current best solution  $s_R^*$  to escape from the current local optimum and potentially explore a new region of the search space (see §4.1.2). This produces a new start point  $s'_R$  for the VND which computes a new local optimum  $s_R$  to the E-VRP-NL (see §4.1.1) and also returns the best solution  $s_O$  to the E-VRP-NL-C it has encountered. Note that we only check the capacity constraints after accepting a move (i.e., when building a new solution to the E-VRP-NL). If appropriate, the algorithm updates the best-known solutions  $s_R^*$  and  $s_O^*$ , as well as the sets  $\Phi_R$  and  $\Phi_O$  of improving solutions. It also populates a pool of routes  $\Omega$ . Note that we only add a route to  $\Omega$  if it does not already contains a route visiting the same sequence of nodes. This procedure is reiterated until the targeted number of iterations has been reached. The optimization returns the sets of improving solutions to the E-VRP-NL and E-VRP-NL-C and the generated pool of routes  $\Omega$ . To speed up the ILS algorithm, its implementation is based on the static move descriptor (SMD) concept introduced by Zachariadis and Kiranoudis (2010) which prevents unnecessary reevaluations of moves and provides an efficient way of exploring neighborhoods (see Appendix B).

---

### Algorithm 2: The ILS algorithm

---

**Input** : a solution  $s_R^0$  to the E-VRP-NL, a solution  $s_O^0$  to the E-VRP-NL-C (possibly equal to NULL), a maximum number of iterations  $\delta^{max}$ , and a maximum route duration limit  $T$

**Output**: a set  $\Phi_R$  and a set  $\Phi_O$  of improving solutions to the E-VRP-NL and to the E-VRP-NL-C, a pool of routes  $\Omega$

```

1 Procedure generateRoutes( $s_R^0, s_O^0, \delta^{max}, T$ ):
   /* We denote  $f(s)$  the value of the objective function for a solution  $s$  and we assume  $f(NULL) = +\infty$  */
2    $\delta \leftarrow 0, \Phi_R \leftarrow \emptyset, \Phi_O \leftarrow \emptyset, \Omega \leftarrow \emptyset$ 
3    $s_R^* \leftarrow s_R^0, s_O^* \leftarrow s_O^0$ 
4   while  $\delta < \delta^{max}$  do
5     if  $\delta = 0$  then  $s'_R \leftarrow s_R^0$ 
6     else  $s'_R \leftarrow \text{perturb}(s_R^*, T)$  (see §4.1.2)
7      $(s_R, s_O) \leftarrow \text{VND}(s'_R, T)$  (see Algorithm 5)
8     addRoutesToPool( $\Omega, s_R$ )
9     if  $f(s_R) < f(s_R^*)$  then  $\Phi_R \leftarrow \Phi_R \cup \{s_R\}, s_R^* \leftarrow s_R$ 
10    if  $f(s_O) < f(s_O^*)$  then  $\Phi_O \leftarrow \Phi_O \cup \{s_O\}, s_O^* \leftarrow s_O$ 
11     $\delta \leftarrow \delta + 1$ 
12  end
13  return ( $\Phi_R, \Phi_O, \Omega$ )

```

---

#### 4.1.1 The VND search phase

The VND relies on an ordered list of LS operators. A single operator is considered at a time. If an improving move is found, the search restarts with the first operator of the list. Otherwise, it moves to the next operator. The search

reaches a local optimum when the last operator fails to improve the current solution.

Our VND employs several classical VRP operators focusing on sequencing decisions. These operators are defined for solutions represented as sequences of customer visits without CSs. We use five vertex exchanges operators that work by relocating or exchanging customer visits : 1-0, 2-0, 1-1, 2-1, and 2-2 vertex exchanges. We also use the inter-route and intra-route versions of 2-opt. We also define a specific operator for the E-VRP-NL, referred to as *separate*. This operator creates two routes from a single route by inserting a return to the depot after a customer visit. It may improve the cost of a solution if at least one CS is part of the split route. Indeed, creating two routes rather than one may decrease the total time when an expensive detour to a CS is avoided. We stop applying an operator as soon as it has found an improving move. We refer to Algorithm 5 in Appendix C for a description of the general scheme of the VND search phase.

We only consider CSs when evaluating LS moves. In order to make charging decisions in such a way that every route involved in a move has the lowest possible duration, we solve one (inter-route moves) or more (intra-route moves) FRVCPs. To this end, we use the labeling algorithm described by Froger et al. (2019). After each call, the duration of the route is stored in a cache memory to avoid recomputing it. Moreover, since only improving solutions are accepted in the VND framework, we only solve FRVCPs for potentially improving moves. Therefore, we rely on a procedure that filters unpromising or infeasible moves. Such moves are determined by establishing a lower bound on the duration of the routes resulting from a move. The value of this lower bound is computed as the sum of two terms: the minimum increase in time to detour to a CS between two successive nodes in the route, and a lower bound on the charging time. The latter term is computed by dividing the sum of the energy consumption of the route and the minimum increase in energy consumption to detour to a CS between two successive nodes in the route, by the steepest slope for a segment of the piecewise linear charging functions. We note that this lower bound corresponds to the true duration of a route in the case where its energy consumption does not exceed  $Q$ . If the lower bound on the route duration exceeds the limit, then the route is infeasible. Otherwise, the procedure checks whether the route duration is stored in a cache memory and modifies the lower bound value accordingly. We use the lower bound on the duration of the routes to determine whether the move is strictly non-improving. We refer to Algorithm 6 in Appendix C for a detailed explanation on this procedure. If a move is not discarded by the lower bound on the basis that it is infeasible or non-improving, we evaluate it exactly when necessary (i.e., before accepting it). To this end, we solve FRVCPs for all the routes that have not been evaluated exactly, as long as the move remains feasible and potentially improving.

#### 4.1.2 The perturbation phase

Whenever we reach a local optimum, we perturb the current solution by removing geographically close customers and by reinserting them at different positions. First, we randomly select a customer  $i \in I$ . We then remove the  $\kappa$  closest customers to  $i$  from their respective routes, with  $\kappa$  randomly selected in the interval  $[\min\{|I|, 5\}, \max\{\min\{|I|, 5\}, \lceil \sqrt{|I|} \rceil\}]$ . We set the distance between two customers  $i_1$  and  $i_2$  as equal to  $0.5 (t_{i_1, i_2} + t_{i_2, i_1} + (e_{i_1, i_2} + e_{i_2, i_1})/\rho^*)$ . The value of  $\rho^*$  corresponds to the steepest slope for a segment of the piecewise linear charging functions. Finally, customers are reinserted in the solution one at a time and in a random order by applying the following rules. A removed customer cannot be reinserted in the same route from which it was removed. We evaluate the increase in time of every feasible insertion of the customer. This is done by reoptimizing the charging decisions. The probability of selecting a given feasible insertion is set inversely proportional to the time increase due to the insertion. If there exists no feasible insertion, we simply create a new route with the customer.

## 4.2 Solution assembler: a branch-and-cut method

The objective of the second component of the matheuristic is to construct the best possible solution to the E-VRP-NL-C from a pool of routes  $\Omega$ , obtained from the route generator component. We recall that the charging decisions made for a route are such that its total duration is minimized. The main challenge of the second component of the heuristic is to combine the routes in a solution while satisfying the capacity constraints.

Deriving the best possible solution by solving a set partitioning (SP) model over a pool of routes is a strategy that has been successfully applied to several hard VRPs (Alvarenga et al., 2007; Subramanian et al., 2013; Villegas et al., 2013; Montoya et al., 2017; Andelmin and Bartolini, 2019). It is used either as a post-optimization phase or as an

intensification phase within a metaheuristic. An example of the latter is the matheuristic proposed by Subramanian et al. (2013) to solve a class of VRPs. This approach has been mostly applied to problems without route coupling constraints (i.e., the feasibility of one route is independent of the feasibility of other routes). Due to the CS capacity constraints, the route coupling constraints need to be accounted for in the E-VRP-NL-C. To the best of our knowledge, only two studies have dealt with route coupling constraints in the assembly phase of a solution from a pool of routes: Morais et al. (2014) and Grangier et al. (2017), both in the context of cross-dock VRPs.

We propose assembling the solutions using a decomposition of the problem into a route selection master problem and a CS capacity management subproblem. The master problem consists in selecting a set of routes such that every customer is covered by exactly one route. The charging decisions within the selected routes (as imported from the first component) may lead to a violation of the capacity constraints. In such cases, the subproblem checks whether the CS capacity constraints can be met by revising some of the charging decisions.

We propose four versions of the subproblem, which primarily depend on the degree of allowed modifications on the selected routes by the master problem. In the first version (denoted by N), we do not revise the charging decisions in the selected routes and we only check whether the CS capacity constraints are satisfied. In the second version (denoted by D), we allow delaying the starting time of each route (i.e., we postpone their starting time) to satisfy the CS capacity constraints. We note that in versions N and D no increase in the total time of the solution is incurred. In the third version (denoted by DW), we allow delaying the starting time of each route, and we also allow vehicles to wait for a charger if a CS is overcrowded by EVs. In DW, waiting times can only occur when multiple CSs are visited within a route, since, when only on CS is visited within a route, delaying the starting time of the route is preferable, as this does not penalize the objective function. In the fourth version (denoted by DWR), we also revise the amount of energy charged at the CSs. If a route contains at least two CSs, we may decide to charge more at the first CS in order to avoid waiting before charging the EV at the second CS. We note that this strategy may sometimes avoid detouring to one of the CSs visited in the original route.

We designed a branch-and-cut method to efficiently solve the problem while exploiting the above discussed decomposition. While solving the route selection problem, we dynamically solve the CS capacity management subproblem. More specifically, at each node, we solve the subproblem for the current selection of routes. Depending on the version of the subproblem, we generate different cuts to discard infeasible route selections. For versions DW and DWR, we also add cuts to account for the underestimation of the increase in the total time to visit the customers.

In the following subsections, we provide a detailed description of our four versions of the CS capacity management subproblem. We use the following notation. The binary parameter  $a_{ri}$  is equal to one if and only if route  $r \in \Omega$  serves customer  $i \in I$ . We define parameter  $t_r$  as the duration of a route  $r \in \Omega$ , obtained from the route generation component. We denote by  $O(\Omega)$  the set containing all the charging operations occurring in the routes belonging to  $\Omega$ , and by  $O_j(\Omega) \subseteq O(\Omega)$  all the charging operations occurring at CS  $j \in F$  in these routes. We denote by  $O(\{r\})$  the list of charging operations occurring in route  $r$ . Let  $r(o)$  and  $j(o)$  be the route and the CS associated with a charging operation  $o$ . Let  $\bar{S}(o)$  and  $\bar{\Delta}(o)$  be the starting time and duration of charging operation  $o$  in the route  $r(o) \in \Omega$ . For each route  $r \in \Omega$ , we assume that the operations in  $O(\{r\})$  are sorted in non-decreasing order of their starting times. We denote by  $\pi(o)$  the charging operation following  $o$  in route  $r(o)$ . If there does not exist any charging operation after  $o$ , we set  $\pi(o) = -1$ . We also denote by  $\Pi^+(o)$  the charging operations occurring after  $o$  in  $r(o)$ . For every route  $r \in \Omega$  and for each charging operation  $o \in O(\{r\})$  we denote by  $t^+(o)$  the total time spent by the EV in  $r$  between the CS associated with  $o$  and the CS associated with  $\pi^+(o)$  or the depot. For each operation  $o$ , we also introduce two parameters  $ES(o)$  and  $LS(o)$  representing its earliest and latest possible starting times. The values of these parameters depend on the different versions of the subproblem and they are specified in the following subsections.

#### 4.2.1 Version N and version D of the subproblem

To model the route selection problem, we introduce a binary variable  $x_r$  equal to 1 if and only if route  $r \in \Omega$  is selected. A MILP formulation of this problem is then the following classical SP model:

$$[HC1] \quad \text{minimize} \quad \sum_{r \in \Omega} t_r x_r \quad (88)$$

$$\text{subject to} \quad \sum_{r \in \Omega} a_{ri} x_r = 1 \quad i \in I \quad (89)$$



$$x_r \in \{0, 1\} \quad r \in \Omega. \quad (90)$$

The objective (88) is to select a subset of routes from  $\Omega$  that minimizes the total duration. Constraints (89) ensure that each customer is visited exactly once. Constraints (90) set the domains of the decision variables.

Let  $\Omega_{1+}(\bar{x})$  denote the set of routes with at least one charging operation resulting from a feasible solution  $\bar{x}$  to the above problem, i.e.,  $\Omega_{1+}(\bar{x}) = \{r \in \Omega : \bar{x}_r = 1 \wedge O(\{r\}) \neq \emptyset\}$ . Note that the routes without mid-route charging need not be considered when checking the capacity constraints. We also define  $\Omega_{1+}(\bar{x}, j) \subseteq \Omega_{1+}(\bar{x})$  as the set of routes with a charging operation at  $j \in F$ .

**Version N** In this version of the subproblem, the CS capacity management subproblem does not revise the charging operations scheduled in the selected routes. It only checks the feasibility of the charging operations with respect to the capacity constraints. We therefore set  $ES(o) = LS(o) = \bar{S}(o)$ . Let  $\text{CMP}_1(\bar{x})$  be this subproblem defined for the routes that belong to  $\Omega_{1+}(\bar{x})$ . It can be decomposed into  $|F|$  independent problems, one for every CS. To solve  $\text{CMP}_1(\bar{x})$ , for every CS  $j \in F$  we apply a polynomial algorithm to check the existence of subsets of operations overloading the CS. Specifically, we call procedure  $\text{CheckCapacityCut}(O_j(\Omega_{1+}(\bar{x})), C_j)$  to check whether the capacity constraints are satisfied according to the scheduled charging operations in  $O_j(\Omega_{1+}(\bar{x}))$ . We refer the reader to Algorithm 7 in Appendix D for the details of this procedure.

**Version D** The main drawback of version N is that it may discard promising routes. Indeed, in some cases simply delaying the starting time of the routes renders them feasible. Note that shifting the start time of a route does not increase the objective function. In version D, we seek to resolve capacity conflicts by shifting starting times of the routes. Let  $\text{CMP}_2(\bar{x})$  be this subproblem. In contrast to  $\text{CMP}_1(\bar{x})$ ,  $\text{CMP}_2(\bar{x})$  does not decompose into an independent problem for each CS. Let  $o$  be a charging operation. Its earliest starting time  $ES(o)$  is equal to  $\bar{S}(o)$  since by construction the operations are left shifted in each route of the pool. The parameter  $LS(o)$  is computed by subtracting from  $T_{max}$  the time needed to complete the route (considering the duration of the next charging operations, the driving times, and no waiting times). Specifically,  $LS(o) = T_{max} - \sum_{o' \in \{o\} \cup \Pi^+(o)} (t^+(o') + \bar{\Delta}(o'))$ .

We now define a continuous-time MILP formulation of  $\text{CMP}_2(\bar{x})$ . Let the variable  $S_o$  define the starting time of operation  $o$ . We model the capacity constraints using a flow-based formulation. For each CS  $j \in F$ , we consider two dummy operations  $\varepsilon_j^+$  and  $\varepsilon_j^-$ , and we define  $\tilde{C}_o := 1$  for all  $o \in O_j(\Omega_{1+}(\bar{x}))$  and  $\tilde{C}_{\varepsilon_j^+} := \tilde{C}_{\varepsilon_j^-} := C_j$ . We then introduce the continuous variable  $f_{oo'}$  representing the quantity of resource (i.e., chargers) that is transferred from charging operation  $o$  to charging operation  $o'$ . We also define the sequential binary variable  $u_{oo'}$  taking the value of 1 if operation  $o$  is processed before operation  $o'$ . The MILP formulation of  $\text{CMP}_2(\bar{x})$  is as follows:

$$[\text{CMP}_2(\bar{x})] \quad \text{minimize } 0 \quad (91)$$

$$\text{subject to } u_{oo'} + u_{o'o} \leq 1 \quad j \in F, o, o' \in O_j(\Omega_{1+}(\bar{x})) : o < o' \quad (92)$$

$$u_{oo''} \geq u_{oo'} + u_{o'o''} - 1 \quad j \in F, o, o', o'' \in O_j(\Omega_{1+}(\bar{x})) \quad (93)$$

$$S_{o'} \geq S_o + \bar{\Delta}(o)u_{oo'} + (LS(o) - ES_{o'})(u_{oo'} - 1) \quad j \in F, o, o' \in O_j(\Omega_{1+}(\bar{x})) \quad (94)$$

$$S_{\pi(o)} = S_o + \bar{\Delta}(o) + t^+(o) \quad o \in O(\Omega_{1+}(\bar{x})) : \pi(o) \neq -1 \quad (95)$$

$$\sum_{o' \in O_j(\Omega_{1+}(\bar{x})) \cup \{\varepsilon_j^+\}} f_{o'o} = 1 \quad j \in F, o \in O_j(\Omega_{1+}(\bar{x})) \quad (96)$$

$$\sum_{o' \in O_j(\Omega_{1+}(\bar{x})) \cup \{\varepsilon_j^+\}} f_{o'o} - \sum_{o' \in O_j(\Omega_{1+}(\bar{x})) \cup \{\varepsilon_j^-\}} f_{oo'} = 0 \quad j \in F, o \in O_j(\Omega_{1+}(\bar{x})) \quad (97)$$

$$\sum_{o \in O_j(\Omega_{1+}(\bar{x})) \cup \{\varepsilon_j^-\}} f_{\varepsilon_j^+, o} = C_j \quad j \in F \quad (98)$$

$$\sum_{o \in O_j(\Omega_{1+}(\bar{x})) \cup \{\varepsilon_j^+\}} f_{o, \varepsilon_j^-} = C_j \quad j \in F \quad (99)$$

$$f_{oo'} \leq \max(\tilde{C}_o, \tilde{C}_{o'}) u_{oo'} \quad j \in F, o, o' \in O_j(\Omega_{1+}(\bar{x})) \quad (100)$$

$$ES(o) \leq S_o \leq LS(o) \quad o \in O(\Omega_{1+}(\bar{x})) \quad (101)$$

$$f_{oo'} \geq 0 \quad j \in F, \quad (102)$$

$$(o, o') \in (O_j(\Omega_{1+}(\bar{x})) \cup \{\varepsilon_j^+\}) \cup (O_j(\Omega_{1+}(\bar{x})) \cup \{\varepsilon_j^-\}) \quad (102)$$

$$u_{oo'} \in \{0, 1\} \quad j \in F, o, o' \in O_j(\Omega_{1+}(\bar{x})). \quad (103)$$

[CMP<sub>2</sub>( $\bar{x}$ )] is a feasibility problem. Constraints (92) state that for two distinct operations  $o$  and  $o'$ , either  $o$  precedes  $o'$ ,  $o'$  precedes  $o$ , or  $o$  and  $o'$  are processed in parallel (if there is more than one charger at the CS). Constraints (93) express the transitivity of the precedence relations. Constraints (94) are the disjunctive constraints on the operations related to the same CS. Each such constraint is active when  $u_{oo'} = 1$  and, in which case, it enforces the precedence relation between charging operations  $o$  and  $o'$ . Note that no waiting times can occur before a charging operation. Constraints (95) enforce the precedence relation and the time lag between the charging operations occurring in the same route. Constraints (96) state that a charger has to be allocated to each charging operation. Constraints (97) ensure flow conservation. Constraints (98) and (99) define the value of the flow leaving the source and the flow entering the sink. Constraints (100) couple the flow variables with the sequence variables. Constraints (101) and (103) define the domains of the decision variables.

#### 4.2.2 Version DW and version DWR of the subproblem

In these versions of the subproblem, we allow a possible increase in the total duration of the routes. Indeed, introducing waiting times at CSs or revising the amount of charged energy may help resolve capacity violations, but such modifications may extend routes due to the non-linearity of the charging functions and the consideration of multiple charging technologies. Let  $\theta$  be a non-negative variable estimating the added delay when solving the CS capacity management subproblem. A MILP formulation of the route selection problem (derived directly from [HC1]) follows:

$$[HC2] \quad \text{minimize} \quad \sum_{r \in \Omega} t_r x_r + \theta \quad (104)$$

$$\text{subject to} \quad (89), (90)$$

$$\theta \geq 0. \quad (105)$$

Thereafter, we assume that we have a fixed selection  $\Omega_{1+}(\bar{x})$  of routes given by fixing the variables  $\{x_r\}_{r \in \Omega}$  to values respecting the current constraints of the route selection problem.

**Version DW** In this version of the subproblem, we assume that EVs can wait at CSs if delaying the starting times of the routes is not sufficient to avoid capacity violations. Let CMP<sub>3</sub>( $\bar{x}$ ) be the scheduling subproblem of the routes  $\Omega_{1+}(\bar{x})$ , which has the objective of minimising the addition of waiting times. The MILP formulation of CMP<sub>3</sub>( $\bar{x}$ ) uses the decision variables  $S_o, f_{oo'}, u_{oo'}$  defined in CMP<sub>2</sub>( $\bar{x}$ ). We also introduce variable  $\nabla_o$  that represents the waiting time incurred before the start of charging operation  $o \in O(\Omega_{1+}(\bar{x}))$ . For every charging operation  $o$ , its earliest starting time  $ES(o)$  is equal to  $\bar{S}(o)$  and its latest starting time  $LS(o)$  is equal to  $T_{max} - \sum_{o' \in \{o\} \cup \Pi^+(o)} (t^+(o') + \bar{\Delta}(o'))$ . The MILP formulation of CMP<sub>3</sub>( $\bar{x}$ ) is as follows:

$$[CMP_3(\bar{x})] \quad \text{minimize} \quad \sum_{o \in O(\Omega_{1+}(\bar{x}))} \nabla_o \quad (106)$$

$$\text{subject to} \quad (92) - (94), (96) - (103)$$

$$S_{\pi(o)} = S_o + \bar{\Delta}(o) + t^+(o) + \nabla_{\pi(o)} \quad o \in O(\Omega_{1+}(\bar{x})) : \pi(o) \neq -1 \quad (107)$$

$$\nabla_o \geq 0, \quad o \in O(\Omega_{1+}(\bar{x})). \quad (108)$$

The objective (106) is to minimize the waiting time inserted in each route. Constraints (107) define the minimum time lag between the charging operations occurring in the same route. Constraints (108) define the domains of the waiting decision variables.

**Version DWR** In this version of the subproblem, in addition to the introduction of waiting times, resolving conflicts at CSs can also be achieved by revising the amounts of energy charged at each CS in every route. We denote by CMP<sub>4</sub>( $\bar{x}$ ) the subproblem in which we want to minimize the increase in the duration of the selected routes. Indeed, revising the

charging operations leads to an increase in time when the CSs in a route have different charging technologies or when the charging functions are non-linear. Note that if we substantially increase the charging amounts at a CS we may not need to visit the subsequent CS in the route. We therefore need to account for the potential removal of visits to CSs.

We denote by  $e^-(o)$  the energy consumption of the EV between its departure from  $i_o$  and its arrival at  $j_{\pi^+(o)}$  if  $\pi(o) \neq -1$ . This takes into account the energy consumed to visit all the customers scheduled in the route between charging operations  $o$  and  $\pi^+(o)$  or the depot. Similarly, we denote by  $e^+(o)$  the energy consumption of the EV from its departure from  $j_{\pi^-(o)}$ , if  $o$  is not the first charging operation of the route or from the depot to its arrival at  $j(o)$ . Since a charging operation can be skipped by charging more energy during the previous or next charging operations of the same route, we define  $\tilde{t}(o)$  and  $\tilde{e}(o)$  as the time and energy saved if the EV does not detour to perform the charging operation  $o$ . We define  $\Omega_{2^+}(\bar{x})$  as the subset of  $\Omega_{1^+}(\bar{x})$  that contains only the routes including at least two charging operations (i.e.,  $\Omega_{2^+}(\bar{x}) = \{r \in \Omega_{1^+}(\bar{x}) : |O(\{r\})| \geq 2\}$ ).

Our formulation of  $\text{CMP}_4(\bar{x})$  draws upon formulation  $[\text{CMP}_3(\bar{x})]$ . Aside from using decision variables defined in the latter,  $[\text{CMP}_4(\bar{x})]$  also uses the following decision variables for the operations in  $O(\Omega_{2^+}(\bar{x}))$ . The variables  $t_j$  and  $\bar{t}_j$  are the scaled arrival and departure times, according to the charging function of CS  $j(o)$ . The binary variables  $\underline{w}_{ok}$  and  $\bar{w}_{ok}$  are equal to 1 if and only if the SoC lies between  $q_{j(o),k-1}$  and  $q_{j(o),k}$ , with  $k \in B_{j(o)} \setminus \{0\}$ , upon starting and finishing operation  $o$ , respectively. The continuous variables  $\underline{\lambda}_{ok}$  and  $\bar{\lambda}_{jk}$  are the coefficients associated with the breakpoints  $(a_{j(o),k}, q_{j(o),k})$  of  $\phi_{j(o)}$  upon starting and finishing operation  $o$ , respectively. The continuous variables  $\underline{y}_o$  and  $\bar{y}_o$  represent the SoC of the EV upon starting and finishing charging operation  $o$ . The continuous variable  $\Delta_o$  represents the duration of charging operation  $o$ . For each route, we check whether it might be possible for a charging operation  $o$  to be skipped by considering that the EV leaves the previous CS (or depot) with a fully replenished battery. If this allows the EV to reach the next CS or to return to the depot without performing  $o$ , then we allow the EV not to detour to the corresponding CS. To this end, we introduce the binary variable  $z_o$  equal to 1 if and only if the charging operation  $o$  is executed. We also compute for every charging operation the time windows during which it must be scheduled. The earliest starting time  $ES(o)$  of a charging operation  $o$  is computed assuming that the EV skips (if the previous computation has shown it is possible) the previous charging operations (if any), and charges the maximum between the energy needed to recover the detour to the CS and the energy required to reach the next CS. To compute the latter, we consider that the SoC of the EV upon arriving at the CS is maximal (a full charge occurs at the previous CS). Then, we estimate the charging times assuming that the EV arrives with an empty battery. The latest starting time  $LS(o)$  of operation  $o$  is computed assuming that the EV returns to the depot at time  $T_{max}$  and assuming that the EV skips the next charging operations (if possible). The MILP formulation of  $\text{CMP}_4(\bar{x})$  is as follows:

$$[\text{CMP}_4(\bar{x})] \text{ minimize } \sum_{o \in O(\Omega_{1^+}(\bar{x}))} \nabla_o + \sum_{o \in O(\Omega_{2^+}(\bar{x}))} (\Delta_o - \bar{\Delta}(o) - (1 - z_o)\tilde{t}(o)) \quad (109)$$

subject to (92) – (93), (97) – (103)

$$\underline{y}_o = \sum_{k \in B_{j(o)}} \underline{\lambda}_{jk} q_{j(o)k} \quad o \in O(\Omega_{2^+}(\bar{x})) \quad (110)$$

$$\underline{t}_o = \sum_{k \in B_{j(o)}} \underline{\lambda}_{jk} a_{j(o)k} \quad o \in O(\Omega_{2^+}(\bar{x})) \quad (111)$$

$$\sum_{k \in B_{j(o)}} \underline{\lambda}_{jk} = \sum_{k \in B_{j(o)} \setminus \{0\}} \underline{w}_{ok} \quad o \in O(\Omega_{2^+}(\bar{x})) \quad (112)$$

$$\sum_{k \in B_{j(o)} \setminus \{0\}} \underline{w}_{ok} = z_o \quad o \in O(\Omega_{2^+}(\bar{x})) \quad (113)$$

$$\underline{\lambda}_{j0} \leq \underline{w}_{o1} \quad o \in O(\Omega_{2^+}(\bar{x})) \quad (114)$$

$$\underline{\lambda}_{jk} \leq \underline{w}_{ok} + \underline{w}_{o,k+1} \quad o \in O(\Omega_{2^+}(\bar{x})), k \in B_{j(o)} \setminus \{0, b_{j(o)}\} \quad (115)$$

$$\underline{\lambda}_{jb_j} \leq \underline{w}_{ob_{j(o)}} \quad o \in O(\Omega_{2^+}(\bar{x})) \quad (116)$$

$$\bar{y}_o = \sum_{k \in B_{j(o)}} \bar{\lambda}_o q_{j(o)k} \quad o \in O(\Omega_{2^+}(\bar{x})) \quad (117)$$

$$\bar{t}_o = \sum_{k \in B_{j(o)}} \bar{\lambda}_o a_{j(o)k} \quad o \in O(\Omega_{2^+}(\bar{x})) \quad (118)$$

$$\sum_{k \in B_{j(o)}} \lambda_{ok} = \sum_{k \in B_{j(o)} \setminus \{0\}} \bar{w}_{ok} \quad o \in O(\Omega_{2+}(\bar{x})) \quad (119)$$

$$\sum_{k \in B_{j(o)} \setminus \{0\}} \bar{w}_{ok} = 1 \quad o \in O(\Omega_{2+}(\bar{x})) \quad (120)$$

$$\bar{\lambda}_{o0} \leq \bar{w}_{o1} \quad o \in O(\Omega_{2+}(\bar{x})) \quad (121)$$

$$\bar{\lambda}_{ok} \leq \bar{w}_{ok} + \bar{w}_{o,k+1} \quad o \in O(\Omega_{2+}(\bar{x})), k \in B_{j(o)} \setminus \{0, b_{j(o)}\} \quad (122)$$

$$\bar{\lambda}_{ob_{j(o)}} \leq \bar{w}_{ob_{j(o)}} \quad o \in O(\Omega_{2+}(\bar{x})) \quad (123)$$

$$\Delta_o = \bar{t}_o - \underline{t}_o \quad o \in O(\Omega_{2+}(\bar{x})) \quad (124)$$

$$\Delta_o \leq a_{j(o), b_{j(o)}} z_o \quad o \in O(\Omega_{2+}(\bar{x})) \quad (125)$$

$$\underline{y}_{o(r,1)} = \bar{q}^{\text{first}}(r) \quad r \in \Omega_{2+}(\bar{x}) \quad (126)$$

$$\underline{y}_{\pi(o)} = \bar{y}_o - e^+(o) + \tilde{e}(o)(1 - z_o) \quad o \in O(\Omega_{2+}(\bar{x})) : \pi(o) \neq -1 \quad (127)$$

$$\bar{y}_o - e^+(o) + \tilde{e}(o)(1 - z_o) \geq 0 \quad o \in O(\Omega_{2+}(\bar{x})) : \pi(o) = -1 \quad (128)$$

$$S_{o'} \geq S_o + \bar{\Delta}(o) u_{oo'} + (LS(o) - ES_{o'})(u_{oo'} - 1) \quad j \in F, o, o' \in O_j(\Omega_{1+}(\bar{x}) \setminus \Omega_{2+}(\bar{x})) \quad (129)$$

$$S_{o'} \geq S_o + \Delta_o + (LS(o) - ES_{o'})(u_{oo'} - 1) \quad j \in F, o, o' \in O_j(\Omega_{2+}(\bar{x})) \quad (130)$$

$$S_{\pi(o)} = S_o + \bar{\Delta}(o) + t^+(o) + \nabla_{\pi(o)} \quad o \in O(\Omega_{1+}(\bar{x}) \setminus \Omega_{2+}(\bar{x})) : \pi(o) \neq -1 \quad (131)$$

$$S_{\pi(o)} = S_o + \Delta_o + t^+(o) - \tilde{t}(o)(1 - z_o) + \nabla_{\pi(o)} \quad o \in O(\Omega_{2+}(\bar{x})) : \pi(o) \neq -1 \quad (132)$$

$$S_o + \Delta_o + t^+(o) - \tilde{t}(o)(1 - z_o) \leq T_{max} \quad o \in O(\Omega_{2+}(\bar{x})) : \pi(o) = -1 \quad (133)$$

$$\sum_{o' \in O_{j(o)}(\Omega_{1+}(\bar{x}) \setminus \Omega_{2+}(\bar{x})) \cup \{\varepsilon_{j(o)}^+\}} f_{o'o} = 1 \quad o \in O(\Omega_{1+}(\bar{x}) \setminus \Omega_{2+}(\bar{x})) \quad (134)$$

$$\sum_{o' \in O_{j(o)}(\Omega_{1+}(\bar{x}) \setminus \Omega_{2+}(\bar{x})) \cup \{\varepsilon_{j(o)}^+\}} f_{o'o} = z_o \quad o \in O(\Omega_{2+}(\bar{x})) \quad (135)$$

$$ES(o) \leq S_o \leq LS(o) \quad o \in O(\Omega_{1+}(\bar{x})) \quad (136)$$

$$z_o \in \{0, 1\}, \Delta_o \geq 0, 0 \leq \underline{y}_o \leq Q, 0 \leq \bar{y}_o \leq Q \quad o \in O(\Omega_{2+}(\bar{x})) \quad (137)$$

$$\underline{w}_{ok}, \bar{w}_{ok} \in \{0, 1\}, \quad o \in O(\Omega_{2+}(\bar{x})), k \in B_{j(o)} \setminus \{0\} \quad (138)$$

$$\underline{\lambda}_{ok}, \bar{\lambda}_{ok} \in \{0, 1\} \quad o \in O(\Omega_{2+}(\bar{x})), k \in B_{j(o)}. \quad (139)$$

The objective (109) is to minimize the total additional time inserted in each route. Constraints (110)–(124) model the piecewise linear charging functions. Constraints (125) impose a duration equal to 0 for each charging operation that is not executed anymore. For each route  $r \in \Omega_{2+}(\bar{x})$ , constraints (126) define the SoC  $\bar{q}^{\text{first}}(r)$  of the EV upon starting its first charging operation. Constraints (127) couple the SoC of the EV after finishing a charging operation with its SoC when starting the next charging operation occurring in the route. Note that if  $z_o$  is equal to 0, then the SoC  $\underline{y}_o = \bar{y}_o$  still takes into account the energy consumed to detour to CS  $j(o)$ . The energy saved by not visiting this CS is subtracted when computing the SoC at the beginning of the next operation of the route or at the arrival at the depot (see (128)). For each route, constraints (128) force the corresponding EV to have enough SoC at the end of the last charging operation to reach the depot. Constraints (129) and (130) are the disjunctive constraints on the operations related to the same CS. Constraints (131) and (132) define a minimum time lag between the charging operations occurring in the same route. Note that if  $z_o$  is equal to 0, then the starting time  $S_o$  still takes into account the detour to CS  $j(o)$ . The time saved by not visiting this CS is subtracted during the computation of the departure time for the next operation of the route. Constraints (133) limit the route duration. Constraints (134) and (135) assign a charger to each operation that is executed. Constraints (136)–(139) define the domains of the new decision variables.

### 4.2.3 Cut generation procedure

To couple the route selection problem with the CS capacity management subproblem, we generate constraints to cut off infeasible selections of routes in  $\Omega_{1+}(\bar{x})$  and to bound the variable  $\theta$  (for versions DW and DWR). The efficiency of our branch-and-cut method is based on the strength of these cuts. Rather than only cutting the current solution, we investigate a strategy that generates cuts for a large portion of the solution space of the route selection problem. Moreover, we try to add several cuts at a time to speed up the convergence of the algorithm.

We first focus on the cuts that are applicable only to version N. Given a CS  $j$ , let  $\mathcal{O}_j$  be a set containing subsets of operations with a cardinality strictly larger than  $C_j$ , which overlap in time. To discard the current solution  $\bar{x}$  in the route selection problem, we add to [HC1] the following cuts:

$$\sum_{r \in \Omega_{1+}(\bar{x}) : \mathcal{O}(\{r\}) \cap \mathcal{O} \neq \emptyset} x_r \leq C_j \quad j \in F, \mathcal{O} \in \mathcal{O}_j. \quad (140)$$

For the other versions of the subproblem, Algorithm 3 summarizes the procedure used to solve the capacity management subproblem and potentially generating cuts. We first try to detect, before solving  $\text{CMP}(\bar{x})$ , the potential infeasibility of the subproblem. To this end, we apply two algorithms (frequently used in scheduling problems) to detect capacity violations that cannot be solved by the CS capacity management subproblem. These algorithms focus on a single CS and are based on a reasoning rooted in the time windows (derived from the opening hours of the depot) in which every charging operation can be scheduled. The first algorithm focuses on the fixed part of the charging operations (line 3). We call a fixed part of a charging operation  $o$  the time interval (if it exists) between the latest starting time  $LS(o)$  and earliest completion time  $ES(o) + \bar{\Delta}(o)$ . Based on the fixed part of the charging operations, we detect subsets of operations that will necessarily lead to a violation of the capacity constraints and we add cuts to forbid the underlying subset of routes (line 6). Hereafter, we refer to this algorithm as the *fixed part based algorithm*. The details of this procedure are provided in Algorithm 8 (Appendix D). The second algorithm is based on an energy reasoning (line 8). The required energy consumption of a charging operation  $o$  during a time interval  $[t_1, t_2)$  is equal to the minimum duration (possibly equal to 0) for which  $o$  is surely executed within the interval. For a CS  $j$ , the total required energy consumption by the charging operation  $o$  scheduled at  $j$  over time interval  $[t_1, t_2)$  cannot exceed  $C_j(t_2 - t_1)$ . The difference between the last term and the total required energy consumption is called the slack over  $[t_1, t_2)$ . The slack over any time interval must always be non-negative. The number of intervals that needs consideration is bounded (see Baptiste et al. (2001) for more details). When there exists an interval for which a slack is strictly negative, the subset of operations having a non-zero energy consumption on this interval cannot be performed without leading to a violation of the capacity constraints. We add a no-good cut to discard it (line 10). Hereafter, we refer to this algorithm as the *energy reasoning based algorithm*. The details of this procedure are provided in Algorithm 9 (Appendix D).

To limit the computation time, if the two previous algorithms prove the infeasibility of the subproblem, we add to the route selection problem the generated cuts and we do not solve the subproblem. Otherwise, since the absence of capacity violation detection from these latter algorithms does not ensure the feasibility of the subproblem, we solve the latter exactly. We observe that the subproblem ( $\text{CMP}_2(\bar{x})$ ,  $\text{CMP}_3(\bar{x})$ , or  $\text{CMP}_4(\bar{x})$ ) may often be decomposed into several independents smaller subproblems. This allows us to potentially formulate cuts for each of these small subproblems. We introduce a graph  $G(\bar{x})$  in which each node represents a CS and there exists an edge between two CSs if there exists a route in  $\Omega_{1+}(\bar{x})$  with charging operations at these two CSs. We can then decompose the subproblem into as many independent subproblems as the number of connected components of  $G(\bar{x})$ . Let  $\mathcal{C}(G(\bar{x}))$  be the connected components of  $G(\bar{x})$ . For every connected component  $C \in \mathcal{C}(G(\bar{x}))$ , only the charging operations scheduled at the CSs in  $C$  need consideration. We denote  $\text{CMP}(\bar{x}, C)$  the subproblem restricted to the CSs in connected component  $C$ . When  $\text{CMP}(\bar{x}, C)$  is infeasible, we generate an integer Benders cut, also called combinatorial Benders cut (Codato and Fischetti, 2006), to invalidate the current solution to the route selection problem (line 18). When it is feasible, we compare the current value  $\bar{\theta}$  of  $\theta$  with the subproblem objective function value. This only holds for versions DW and DWR of the subproblem. To build strong cuts, we consider the set  $\mathcal{C}^{\text{feas}}$  of connected components for which the corresponding subproblem is feasible. For each subset  $C$  of  $\mathcal{C}^{\text{feas}}$ , we generate an integer Benders optimality cut if the route selection problem underestimates the increase in time needed to resolve conflict(s) at CS(s) (line 25).

#### 4.2.4 The branch-and-cut algorithm

We now outline the general structure of the solution assembler component in Algorithm 4. When provided as an input, we give to the solver the objective function value of a solution to the E-VRP-NL-C as a cutoff value (i.e., an upper bound on the value of the objective function of the master problem). Therefore, the solver only considers the solutions with an objective function value strictly less than this cutoff value. The solver returns the best solution assembled from a given pool of routes by the branch-and-cut method. The component also returns a set  $L$  of improving solutions (compared to the initial solution  $s$ ) computed throughout the algorithm.

---

**Algorithm 3:** Solving the CS capacity management subproblem and generating cuts for version D, DW, or DWR

---

**Input:** a solution  $\bar{x}$  to the current master problem

```

1 Procedure SolveSubProblem( $\bar{x}$ ):
2   Compute  $ES(o)$  and  $LS(o)$  for each charging operation  $o \in O(\Omega_{1+}(\bar{x}))$ 
3   for  $j \in F$  do
4      $\mathcal{U} \leftarrow \text{FixedPartsAlgorithm}(O_j(\bar{x}), C_j)$  (see Algorithm 8)
5     if  $\mathcal{U} \neq \emptyset$  then
6       for  $U \in \mathcal{U}$  do Generate the following cut and add it to  $[HC]$ :  $\sum_{r \in \Omega_{1+}(\bar{x}, j) : O(\{r\}) \cap U \neq \emptyset} x_r \leq C_j$ 
7     else
8        $\mathcal{U} \leftarrow \text{EnergeticAlgorithm}(O_j(\bar{x}), C_j)$  (see Algorithm 9)
9       if  $\mathcal{U} \neq \emptyset$  then
10        for  $U \in \mathcal{U}$  do Generate the following cut and add it to  $[HC]$ :
11           $\sum_{r \in \Omega_{1+}(\bar{x}, j) : O(\{r\}) \cap U \neq \emptyset} x_r \leq |\{r \in \Omega_{1+}(\bar{x}, j) : O(\{r\}) \cap U \neq \emptyset\}| - 1$ 
12        end
13      end
14    end
15  if no cuts have been generated and  $\bar{x}$  is integer then
16     $\mathcal{C}^{\text{feas}} \leftarrow \emptyset$ 
17    for each  $C \in \mathcal{C}(G(\bar{x}))$  do
18      /* Let  $\text{CMP}(\bar{x}, C)$  be the subproblem restricted to the CSs in connected component  $C$  and  $z(\text{CMP}(\bar{x}, C))$  denotes
19      the value of the objective function (when a solution is found) */
20      Solve the MILP formulation of  $\text{CMP}(\bar{x}, C)$  that corresponds to the selected version of the subproblem
21      if no solution is found then
22        Generate the following cut and add it to  $[HC]$ :  $\sum_{j \in C} \sum_{r \in \Omega_{1+}(\bar{x}, j)} x_r \leq \sum_{j \in C} |\Omega_{1+}(\bar{x}, j)| - 1$ 
23      else
24        /* Only for versions DW and DWR
25         $\mathcal{C}^{\text{feas}} \leftarrow \mathcal{C}^{\text{feas}} \cup \{C\}$ 
26      end
27    end
28    /* Only for versions DW and DWR
29    for every subset  $\mathcal{C}$  of connected components of  $\mathcal{C}^{\text{feas}}$  do
30      if  $\sum_{C \in \mathcal{C}} z(\text{CMP}(\bar{x}, C)) > \bar{\theta}$  then Generate the following cut and add it to  $[HC]$ :
31        
$$\left( \sum_{C \in \mathcal{C}} z(\text{CMP}(\bar{x}, C)) \right) \left( 1 + \sum_{C \in \mathcal{C}} \sum_{j \in C} \sum_{r \in \Omega_{1+}(\bar{x}, j)} x_r - \sum_{C \in \mathcal{C}, j \in C} |\Omega_{1+}(\bar{x}, j)| \right) \leq \theta$$

32      end
33    end
34  if no cuts have been generated then
35    Save the solution result of calling the subproblem on the routes of  $\Omega_{1+}(\bar{x})$ 
36  end

```

---

---

**Algorithm 4:** The branch-and-cut algorithm

---

**Input** : a pool  $\Omega$  of routes / a solution  $s$  (possibly equal to `NULL`) to the E-VRP-NL-C / a CPU time limit  $\tau$  for the algorithm / a CPU time limit for each call to the subproblem  $\tau^{SP}$

**Output:** the best solution  $s^*$  computed by the branch-and-cut method / a set  $L$  of improving solutions to the E-VRP-NL-C (in comparison with  $s$ )

1 **Procedure** `assemble`( $\Omega, s, \tau, \tau^{SP}$ ):

2     Build the route selection problem  $[HC.]$  from pool  $\Omega$

3     Define the callback function (function called at every node of the branch-and-bound tree) as procedure `SolveSubProblem`( $\bar{x}$ ) (see Algorithm 3) and associate to it a CPU time limit of  $\tau^{SP}$  per call

4     Define the cutoff value  $f(s)$  (objective function value of  $s$ ) and the CPU time limit  $\tau$  of the solver

5     Give the model  $[HC.]$  and the callback function to the solver and launch it

6     Retrieve the best solution  $s^*$  computed within the CPU time limit and the set  $\Phi$  of improving solutions (compared to  $s$ )

7     **return** ( $s^*, \Phi$ )

---

## 5 Computational results

We used Gurobi 7.5.0 (through its Java API) to solve the MILP models. Each instance was executed on a single thread with 12 GB and on a cluster of 27 computers, each having 12 cores and two Intel(R) Xeon® X5675 3.07 GHz processors. We implemented all the algorithms using Java 8. In all the tables, the CPU time is given in seconds and rounded to the nearest integer.

We considered the 120-instance testbed of Montoya et al. (2017)<sup>1</sup>. In this testbed, there are six sets of 20 instances, each with 10, 20, 40, 80, 160, or 320 customers. The EVs are Peugeot iOns which have a consumption rate of 0.125 kWh/km and a battery of 16 kWh. When dealing with the E-VRP-NL-C, we adapted each instance by fixing a number of chargers for each CS. We decided to consider instances in which all the CSs have one or two chargers.

The first aim of our computational experiments is to assess and compare the performance of the CS replication-based and the path-based models (Section 3) for solving small-size instances of the E-VRP-NL-C. These results are presented in §5.1. The second aim of our computational experiments is to assess the quality of the ILS presented in §4.1 as a solution method for the E-VRP-NL. A comparison with results from the literature is presented in §5.2. The third aim of our experiments is to assess the effectiveness of our algorithmic framework to build high-quality solutions to the E-VRP-NL-C. We compare the results given by our matheuristic according to the version of the subproblem selected during the assembly phase. These results are presented in §5.3.

### 5.1 Results for E-VRP-NL-C formulations

For the E-VRP-NL-C we considered the twenty 10-customer and the twenty 20-customer instances. We ran the MILP models with a three-hour CPU time limit. In the CS replication-based models  $[F^{CS-rep}]$ , the number of copies of each CS was set to  $\beta \geq 1$ . Since there does not yet exist a procedure for fixing  $\beta$  while guaranteeing optimality, we tested  $\beta = 4$  and  $\beta = 5$ .

Table 1 reports for the CS replication-based and path-based formulations the number of instances with a solution proven to be optimal ( $\#Opt$ ), the average CPU time in seconds (Time) for these latter instances, and the average gap (Gap) for the remaining instances. We compute the gap as  $(z - z^{LB})/z$ , where  $z$  is the objective function value of the best integer solution returned by the solver, and  $z^{LB}$  is the best lower bound retrieved by the solver. The detailed results for all the tested instances are reported in Appendix E.1.

Our results show that the path-based model outperforms the CS replication-based models on the 10-customer instances. Indeed, the MILP solver can optimally solve all the 10-customer instances within an average computation time of roughly 10 minutes for the former model, whereas four instances cannot be solved to optimality within the CPU time limit for the latter model. We also observe that the 20-customer instances are already challenging for the models since very few instances were solved. The introduction of the capacity constraints weakens the dominance rules between CSPs in the path-based model. This is due to the waiting times which are difficult to estimate and embed within the filtering technique. The number of paths therefore grows very quickly. Regarding the CS replication-based model,

---

<sup>1</sup>available from <http://www.vrp-rep.org/> (Retrieved October 25, 2019)

increasing the value of  $\beta$  increases computation times. As expected, since the linear relaxation of each of these models is weak, we conclude that solving the MILP formulations with a commercial solver does not constitute an efficient solution method for the E-VRP-NL-C. Still, 22 out of the 40 tested instances were optimally solved when considering one or two chargers per CS.

Table 1: Computational results for the CS replication-based and path-based formulations on the 10-customer and 20-customer instances.

Capacity	$ I $	$[F^{CSrep}] (\beta = 4)$			$[F^{CSrep}] (\beta = 5)$			$[F^{path}]$		
		#Opt	Time	Gap	#Opt	Time	Gap	#Opt	Time	Gap
1	10	16	372	14.4%	16	503	14.7%	20	572	-
	20	5	1464	12.7%	5	708	12.8%	2	543	20.5%
2	10	16	431	13.6%	16	581	16.3%	20	621	
	20	5	597	*13.2%	5	1387	*14.2%	2	656	21.0%

\*: There are 4 instances for which the solver reaches the CPU time limit without finding any solution. They are not considered when computing the gap.

## 5.2 Results for the E-VRP-NL

Since the route generator of our matheuristic essentially solves the E-VRP-NL, we wanted to assess its quality on this problem. We considered the 120 instances of the original Montoya et al. (2017) testbed. We performed 10 test runs for each instance using Algorithm 2 and compared the results with those obtained in the E-VRP-NL literature through two solutions methods: the metaheuristic of Montoya et al. (2017) which combines an ILS with a heuristic concentration (HC), and the large neighborhood search (LNS) of Koç et al. (2018). Table 2 shows the results of this comparison. Given a number of customers in the instances, it reports for each solution method the number of BKS to the E-VRP-NL (#BKS), the average gap to the BKS (Gap), and the average number of routes in the best computed solution. We compute the gap as  $(z - z^*)/z$ , where  $z$  is the objective function value of the best solution returned by the solution method, and  $z^*$  is the objective function value of the BKS. Table 2 also reports the average gap to the best average objective function value (Gap BA) as  $(z_{avg} - z_{avg}^*)/z_{avg}$ , where  $z_{avg}$  is the average objective function value of the solutions obtained in 10 runs by the solution method, and  $z_{avg}^*$  is the best average (BA) objective function value obtained in 10 runs by one of the existing solution method for the E-VRP-NL. The detailed results for all the tested instances are reported in Appendix E.2.

Table 2: Comparison of the results obtained by ILS with the results of Montoya et al. (2017) and Koç et al. (2018) for the E-VRP-NL.

$ I $	ILS + HC 2.33 GHz processor - 16GB of RAM					LNS 3.6 GHz processor - 32 GB of RAM					ILS 3.07 GHz processor - 12 GB of RAM						
	#BKS	Gap	BKS	#R	Gap BA	Time	#BKS	Gap	BKS	#R	Gap BA	Time	#BKS	Gap	BKS	#R	Gap BA
10	20	0.0%	2.7	0.3%	6	20	0.0%	2.7	0.1%	8	20	0.0%	2.7	0.0%	5		
20	11	0.3%	3.7	0.7%	11	12	0.2%	3.6	0.4%	14	20	0.0%	3.6	0.0%	8		
40	3	1.0%	6.5	2.6%	35	6	0.9%	6.4	1.1%	45	20	0.0%	6.2	0.0%	20		
80	0	3.8%	9.2	5.4%	80	0	3.8%	8.8	3.8%	99	20	0.0%	8.6	0.0%	64		
160	0	7.3%	16.7	8.1%	568	0	7.7%	16.6	7.8%	632	20	0.0%	15.3	0.0%	295		
320	0	11.2%	32.6	12.6%	4398	0	11.7%	32.6	12.4%	4555	20	0.0%	29.0	0.0%	1118		
All	34	3.9%	11.9	4.9%	850	38	4.1%	11.8	4.3%	892	120	0.0%	10.9	0.0%	252		

ILS + HC: (Montoya et al., 2017), LNS: (Koç et al., 2018).

The results presented in Table 2 clearly show that our ILS outperforms all existing methods for the E-VRP-NL. We have identified 80 new BKS and matched the existing BKS for the remaining instances. We have improved the previous solutions by about 4.0%. The improvement increases with the number of customers. The algorithm is also stable as the average results on 10 test runs for each instance are better than those reported for previous methods. While it is



difficult to draw definitive conclusions on the computation times since these tests were run on different machines, it seems that the ILS is at least as fast as the other methods, if not the fastest one. The major difference between our method and the existing ones comes from the reoptimization of the charging decisions when evaluating a move.

### 5.3 Results for the E-VRP-NL-C

We have performed several tests to assess the effectiveness of our matheuristic in obtaining high-quality solutions for the E-VRP-NL-C. After some preliminary experiments, we set the values of the parameters of the matheuristic as presented in Table 3. Setting the stopping criterion to 12 iterations and the number of ILS iterations to 200 proved to be a good compromise between solution quality and computation time. In some rare cases we were not able to optimally solve the assembly phase using version DWR of the subproblem. Indeed, when it is difficult to find a solution satisfying the capacity constraints or when such a solution does not exist, the MILP formulation becomes computationally expensive. Nonetheless, after testing higher CPU time limits for the second component, we observed that the impact on the results was negligible.

Table 3: Value of the matheuristic parameters.

$n_{max}$	$\delta$	$\tau$	$\tau^{SP}$	$\alpha$	$T_{min}$
12	200	180 s	5 s	0.9	$0.67 \cdot T_{max}$

We first compare the results obtained by our matheuristic for the four versions of the subproblem used by the solution assembler. For one or two chargers at each CS and each version of the subproblem, we performed 10 test runs for each instance. Table 4 reports the best results obtained during our tests. Specifically, given a number of chargers per CS and a selected version of the subproblem used in the branch-and-cut method, this table reports the number of instances with a solution in each of the 10 tests (#F), the number of instances with the best solution (BS) to the E-VRP-NL-C computed all over our tests (#BS), including those of the path-based formulation, and the average gap to the BS (Gap BS). Table 5 reports the average CPU time in seconds over 10 runs for the whole algorithm ( $T$ ), and for the first ( $T_1$ ) and second ( $T_2$ ) components. It also reports the average gap with respect to the best average objective function value obtained using a given version of the subproblem (Gap BA). The gaps are computed as in §5.2. Detailed results for all the tested instances are reported in Appendix E.3.

Table 4: Comparison of the best results obtained by the matheuristic for the E-VRP-NL-C according to the version of the subproblem it uses and the number of chargers at each CS.

Cap.	I	Version N			Version D			Version DW			Version DWR		
		#F	#BS	Gap BS	#F	#BS	Gap BS	#F	#BS	Gap BS	#F	#BS	Gap BS
1	10	20	17	0.06%	20	20	0.00%	20	20	0.00%	20	20	0.00%
	20	20	15	0.09%	20	20	0.00%	20	20	0.00%	20	20	0.00%
	40	19	15	0.17%	19	18	0.01%	19	18	0.01%	19	19	0.00%
	80	20	15	0.02%	20	20	0.00%	20	20	0.00%	20	20	0.00%
	160	19	5	0.27%	20	13	0.06%	20	13	0.06%	20	15	0.09%
	320	10	5	0.32%	20	1	0.28%	20	2	0.27%	20	13	0.10%
<b>All</b>		<b>105</b>	<b>72</b>	<b>0.14%</b>	<b>119</b>	<b>92</b>	<b>0.06%</b>	<b>119</b>	<b>93</b>	<b>0.06%</b>	<b>119</b>	<b>107</b>	<b>0.03%</b>
2	10	20	20	0.00%	20	20	0.00%	20	20	0.00%	20	20	0.00%
	20	20	20	0.00%	20	20	0.00%	20	20	0.00%	20	20	0.00%
	40	19	19	0.09%	20	20	0.00%	20	20	0.00%	20	20	0.00%
	80	20	18	0.01%	20	20	0.00%	20	20	0.00%	20	19	0.01%
	160	20	13	0.05%	20	10	0.18%	20	10	0.18%	20	14	0.09%
	320	19	4	0.33%	20	9	0.12%	20	9	0.12%	20	8	0.17%
<b>All</b>		<b>118</b>	<b>94</b>	<b>0.08%</b>	<b>120</b>	<b>99</b>	<b>0.05%</b>	<b>120</b>	<b>99</b>	<b>0.05%</b>	<b>120</b>	<b>101</b>	<b>0.05%</b>

Table 5: Comparison of the average results obtained by the matheuristic for the E-VRP-NL-C according to the version of the subproblem it uses and the number of chargers at each CS.

Cap.	$ I $	Version N				Version D				Version DW				Version DWR			
		$T$	$T_1$	$T_2$	Gap BA	$T$	$T_1$	$T_2$	Gap BA	$T$	$T_1$	$T_2$	Gap BA	$T$	$T_1$	$T_2$	Gap BA
1	10	6	5	1	0.06%	6	5	1	0.00%	6	5	1	0.00%	6	5	1	0.00%
	20	11	10	1	0.09%	11	10	1	0.00%	11	10	1	0.00%	11	10	1	0.00%
	40	24	23	1	0.31%	23	22	1	0.03%	23	22	1	0.03%	87	22	65	0.00%
	80	74	73	1	0.10%	74	73	1	0.00%	73	72	1	0.00%	74	73	1	0.04%
	160	340	335	5	0.45%	339	335	4	0.10%	340	336	4	0.10%	397	333	64	0.08%
	320	1345	1274	71	0.34%	1480	1253	227	0.17%	1487	1259	228	0.15%	1933	1244	689	0.05%
<b>All</b>		<b>300</b>	<b>287</b>	<b>13</b>	<b>0.22%</b>	<b>322</b>	<b>283</b>	<b>39</b>	<b>0.05%</b>	<b>323</b>	<b>284</b>	<b>39</b>	<b>0.05%</b>	<b>418</b>	<b>281</b>	<b>137</b>	<b>0.03%</b>
2	10	5	5	0	0.00%	5	5	0	0.00%	5	5	0	0.00%	5	5	0	0.00%
	20	10	10	0	0.00%	10	10	0	0.00%	10	10	0	0.01%	10	10	0	0.00%
	40	23	23	0	0.05%	23	23	0	0.11%	23	23	1	0.09%	45	23	22	0.01%
	80	73	72	1	0.18%	73	73	1	0.19%	73	73	1	0.25%	73	73	1	0.14%
	160	336	334	2	0.24%	336	334	2	0.25%	337	335	2	0.24%	338	336	2	0.28%
	320	1340	1279	61	0.21%	1283	1276	7	0.12%	1276	1269	7	0.14%	1327	1269	58	0.08%
<b>All</b>		<b>298</b>	<b>287</b>	<b>11</b>	<b>0.11%</b>	<b>288</b>	<b>287</b>	<b>2</b>	<b>0.11%</b>	<b>287</b>	<b>286</b>	<b>2</b>	<b>0.12%</b>	<b>300</b>	<b>286</b>	<b>14</b>	<b>0.09%</b>

First, it should be noted that our matheuristic returns an optimal solution for the 22 small-size instances for which the MILP path formulation yields an optimal solution in about 10% of the computation time. Considering that a single charger exists in each CS, we observe that using version N of the subproblem prevents the algorithm from finding solutions to the E-VRP-NL-C for 15 instances. In this case, the algorithm finds the best solution (computed all over our tests) for only 72 out of 120 instances. Delaying the starting time of the routes yields better solutions for 20 more instances. In contrast, allowing waiting times in version DW leads to very marginal improvements, compared with version D. Indeed, making the EVs wait for an available charger can eliminate capacity violations only if at least two CSs are visited in a route; otherwise it is sufficient to make the EVs leave the depot after time 0, as is done in version D. Although the same requirement (more than one CS) holds when allowing the revision of the amount of energy charged at the CSs in version DWR, the latter strategy yields almost all the best solutions to the E-VRP-NL-C.

When increasing the number of chargers to two per CS, the capacity constraints become less binding and all our assembly strategies yield very similar results. This was somewhat expected but it should be noted that delaying the starting time of the routes increases the probability of ending up with a feasible solution.

The first general conclusion is that allowing delays when solving the CS capacity management subproblem is a suitable and efficient way of assembling high-quality solutions to the E-VRP-NL-C. Allowing the revision of the amounts of energy charged at the CSs on top of it significantly improves the results, but it is slightly more computationally expensive (Table 5). We also note that in general, most of the computation time is spent generating routes as assembling a solution is usually very fast, especially when a cutoff value is provided to the branch-and-cut method. Not surprisingly the computation time increases with the number of customers but it remains reasonable since it takes around 30 minutes to tackle the largest instances.

To assess the relevance of our cuts, we show in Table 6 the distribution of the different type of cuts generated in the branch-and-cut tree. When the capacity constraints are very binding, the energetic reasoning based algorithm detects infeasible route selections much more frequently than the fixed part reasoning based algorithm. We observe that due to its high degree of flexibility, compared with the other versions, DWR requires solving the MILP formulation more often. Indeed, the two previously mentioned algorithms are weaker in that case since the mandatory part of each charging operation may have a very short duration due to the potential revision of the amount of energy charged at the CSs. When we increase the number of chargers at each CS, the reasoning algorithms lead to the generation of fewer cuts. This is not surprising since these are based on relaxations of the subproblem. However, they are useful since the generated cuts are stronger, which avoids computationally expensive calls to the MILP solver. We also see that integer optimality cuts are seldom generated. Since increases in the total time of the routes are penalized, the algorithm tends to avoid adding waiting times. This indicates that revising the charging operations does not always lead to an increase in time,

and a trade-off between delay and a revision at no cost is often found. To support this conclusion, we have analyzed the solutions returned by the matheuristic. Given the number of chargers at each CS, Table 7 reports the percentage of solutions for which there exists at least one route with a delayed starting time, a waiting time before a charging operation, and a revision of the amount of energy charged in the EV. We see that for most of the solutions returned by the algorithm, satisfying the capacity constraints can be achieved without increasing the cost.

Table 6: Characteristics of the cuts generating in the solution assembly phase.

Cap.	Version D			Version DW				Version DWR			
	FP	NRG	BF	FP	NRG	BF	BO	FP	NRG	BF	BO
1	2.3%	94.5%	3.2%	2.3%	95.0%	2.5%	0.2%	0.6%	16.9%	80.5%	2.0%
2	29.0%	28.2%	42.8%	28.9%	27.1%	42.5%	1.5%	1.3%	0.7%	96.2%	1.8%

FP (feasibility cuts generated after calling the fixed part reasoning based algorithm), NRG (feasibility cuts generated after calling the energetic based reasoning algorithm), BF (feasibility cuts of type generated after solving the MILP formulation), BO (optimality cuts of type generated after solving the MILP formulation)

Table 7: Analysis of the solutions retrieved by the matheuristic.

Capacity	Version D	Version DW		Version DWR				
	= D	= D	> W	= D	= NRG	> W	> NRG	> R
1	42.1%	42.2%	1.7%	42.0%	17.4%	1.3%	6.3%	2.5%
2	11.7%	11.7%	0.0%	11.4%	8.5%	0.0%	0.4%	0.0%

D (delay), W (waiting time), NRG (revision of the amount of energy charged), R (removal), = (no increase of the total time), > (increase of the total time)

Since we have imposed a CPU time limit for the branch-and-cut execution and since the maximum route duration limit may decrease according to the state of the solution method, no version of the subproblem dominates the other versions. To perform a fair comparison between the different strategies, and to quantify the benefit of allowing the addition of waiting times and the revision of the amount of energy charged at the CSs, we took the 2,400 long-term pools of routes obtained at the end of the matheuristic for version DWR of the subproblem. For every pool of routes, we ran the branch-and-cut algorithm to generate the best E-VRP-NL-C solution. Table 8 reports for each version of the subproblem the number of pool of routes for which the algorithm obtains a feasible solution (#F), as well as the best solution (#Best) among those obtained with the other versions on the same pool. It also shows the average gap between the solution returned and the best solution obtained during these tests under all the assembly strategies. The conclusions are very similar to those drawn from the results of the matheuristic. Compared to version N, version DWR makes nearly 400 extra routes feasible when assuming one charger. Not surprisingly, when increasing the number of chargers, the results tend to be more independent of the selected version of the subproblem.

Table 8: Comparison of the results of the branch-and-cut method according to the different versions of the subproblem.

Cap.	Version N			Version D			Version DW			Version DWR		
	#F	#Best	Gap	#F	#Best	Gap	#F	#Best	Gap	#F	#Best	Gap
1	1916	1361	0.315%	2262	2095	0.096%	2270	2112	0.091%	2307	2307	0.000%
2	2321	2125	0.073%	2390	2371	0.013%	2390	2371	0.013%	2390	2390	0.000%

## 6 Conclusion and perspectives

We have modeled and solved an electric vehicle routing problem that embeds several features such as piecewise linear charging functions, multiple charging technologies, and multiple visits to CSs. Our methodology explicitly considers the fact that the number of EVs simultaneously charging at every CS is limited by the number of chargers. We have proposed two continuous-time formulations of the E-VRP-NL-C: a CS replication-based and a path-based formulation. Our results show that optimally solving even small-size instances is challenging. To handle larger instances, we have developed an algorithmic framework which iteratively calls a route generator and a solution assembler. The first component focuses on generating a pool of high-quality and diversified routes from solutions to the E-VRP-NL obtained by means of an ILS metaheuristic. Computational experiments have shown that this first component produces high quality E-VRP-NL solutions. Indeed, we have improved 80 out of 120 best-known E-VRP-NL solutions. The second component assembles a solution to the E-VRP-NL-C by selecting a subset of routes from the generated pool. We decomposed this assembly problem into a route selection problem and a CS capacity management subproblem. We have designed a branch-and-cut method based on this decomposition scheme. We have developed and compared four versions of the CS capacity management subproblem, including a simple check of the capacity constraints, the introduction of waiting times, and the revision of the charging amounts. Our results show that there exists a serious risk of ending up with no solution if the method disregards the capacity constraints or if the routes cannot be modified by the subproblem. Delaying the starting time and revising the amount of energy charged at the visited CSs lessens this risk. Using more complex strategies to solve capacity violation issues at CSs does not significantly increase the computation time and yields better solutions. The matheuristic is also able to find all optimal solutions of small-size instances.

*Acknowledgments*— This research was partly funded by the French Agence Nationale de la Recherche through project e-VRO (ANR-15-CE22-0005-01) and by the Canadian Natural Sciences and Engineering Research Council under grants 436014-2013 and 2015-06189. This support is gratefully acknowledged.

## A Experiments on the feasibility of solutions from the literature when considering capacitated CSs

We verified the feasibility of the 120 best-known solutions (BKSs) for the E-VRP-NL reported in Montoya et al. (2017), by limiting the number of chargers per CS to one, two, three, and four. Table 9 presents the results of our experiments. We observe that if there exists only one charger at each CS, then almost half of the solutions are infeasible (i.e., they violate the capacity constraints). This proportion drops to 11% when allowing two chargers per station, and four chargers at each CS are needed to ensure the feasibility of all solutions. In practice, however, there are usually only one or two chargers available at each CS.

Table 9: Results of the feasibility tests performed on the best solutions obtained in (Montoya et al., 2017) (All CSs have the same number of chargers).

Experiment	# Infeasible solutions	Average duration with capacity violation	Average #EVs during capacity violation
$C_1$	55/120	32 min	2.1
$C_2$	23/120	23 min	3.2
$C_3$	3/120	33 min	4.0
$C_4$	0/120	0 min	0.0

## B Implementation details for the ILS

Our implementation of the ILS is based on the static move descriptor (SMD) concept. An SMD is static information that describes a move independently of the current solution. A cost tag is associated with every SMD to store its evaluation and it is updated throughout the LS. By storing SMDs in special data structures, one can access and update the moves in an efficient way in order to reduce computation time. Specifically, SMDs are stored in a priority queue (PQ) and organized according to their dynamic cost tag. The exploration of a neighborhood consists in looping over the PQ until a feasible SMD can be applied. Then, to avoid unnecessary reevaluations of moves, only the cost tag of the SMDs impacted by the previously applied SMD is updated. One major difference with the previous use of the SMD framework lies in the complexity of the evaluation of the moves. To our knowledge, until now there has been an emphasis on feasibility issues when using the SMD framework. Indeed, since checking the feasibility of each move can be time consuming, this is only done during the exploration phase of the neighborhood. In our case, not only the feasibility of a move can be computationally expensive to check, but also its evaluation.

Our SMD implementation is inspired from the one proposed by Beek et al. (2018) who implemented the PQ with a binary heap<sup>2</sup>. For every operator, we use a matrix data structure to store all the SMDs. We build these matrices during the initialization phase of the algorithm and we store them during the entire runtime of the algorithm. To avoid a computationally expensive initialization phase of the ILS, the cost tag of each SMD is set equal to the value computed after the first step of the evaluation (see §4.1.1). Since we are only interested in the moves that can potentially improve the current solution, only those SMDs with a cost tag strictly below 0 need consideration (we refer to them as improving SMDs). Throughout the algorithm, we use two binary heaps to store potentially improving SMDs: an “exact” binary heap (EBH) for the SMDs that are exactly evaluated and a “approximate” binary heap (ABH) for the remaining SMDs. In the initialization of the algorithm, we add to EBH all the improving SMDs that are exactly evaluated using the first step, and we add to ABH the other improving SMDs. When we start exploring a neighborhood, we apply the first move in EBH if it is not empty. Otherwise, we iterate over ABH. We apply the second step of the evaluation to compute the exact cost tag of each SMD in ABH until we find an improving move or we reach the end of the heap. After selecting a move  $m$ , we update the values of the SMDs if they contain at least one node in the routes impacted by

<sup>2</sup>A binary heap is a complete binary tree satisfying the heap ordering property: the value of the key stored in each node is less than or equal to the value of the keys in the node’s children.

*m*. Although in some cases the cost tag of a certain number of these SMDs may remain identical, we want to keep the SMD framework as simple as possible. To avoid a computationally expensive update phase, each cost tag is set equal to the value computed during the first step evaluation. It can therefore be a lower bound on the exact evaluation of the move. In contrast with Beek et al. (2018), we keep cross-operator effects, meaning that we update the cost tag of SMDs associated with other operators. However, we perform this update for the other operators only when the search moves to them. To this end, we associate with each route the iteration number for which has been lastly updated, and to each operator the iteration number for which it has last been called (we increment the iteration number each time the search accepts a move or changes operator).

## C Algorithmic details for the ILS

Algorithm 5 describes the general scheme of the VND search phase. Algorithm 6 describes the move evaluation procedure.

---

### Algorithm 5: The VND algorithm

---

```

Input : a solution  $s^0$  to the E-VRP-NL and a maximum route duration limit  $T$ 
Output: a solution  $s_R$  to the E-VRP-NL and a solution  $s_O$  (possibly equal to NULL) to the E-VRP-NL-C
1 Procedure VND( $s^0, T$ ):
   /* We denote  $f(s)$  the value of the objective function for a solution  $s$  and we assume  $f(\text{NULL}) = +\infty$  */
2    $s_R \leftarrow s^0, s_O \leftarrow \text{NULL}, k \leftarrow 0$ 
3    $N \leftarrow [(1-0, 2-0 \text{ vertex exchanges}), (1-1, 2-1, 2-2 \text{ vertex exchanges}), (2\text{-opt intra and inter-routes}), \text{separate}]$ 
4   while  $k < |N|$  do
     /* The procedure  $\text{search}(N[k], s, T)$  searches an improving solution to  $s$  in the neighborhood  $N[k]$  respecting the
       maximum route duration limit  $T$ . It uses Algorithm 6 for move evaluation */
5      $s' \leftarrow \text{search}(N[k], s_R, T)$  /*  $s' = \text{NULL}$  if no improving solution is found */
6     if  $s'$  is a solution to the E-VRP-NL-C and  $f(s') < f(s_O)$  then  $s_O \leftarrow s'$ 
7     if  $f(s') < f(s_R)$  then  $s_R \leftarrow s', k \leftarrow 0$ 
8     else  $k \leftarrow k + 1$ 
9   end
10  return ( $s_R, s_O$ )

```

---

## D Algorithmic details for the branch-and-cut method

Algorithm 7 describes the procedure to check the capacity constraints for version N of the subproblem. Algorithm 8 and Algorithm 9 provide a detailed description of the fixed part and energy based reasoning algorithms.

---

**Algorithm 6:** Evaluation of a move

---

**Input** : a move  $m$  and a maximum route duration limit  $T$   
**Output**: a couple  $(c,b)$  where  $c$  is a lower estimation of the move evaluation and  $b$  is a boolean equal to **true** if the move is exactly evaluated or if it is a non-improving move and it is equal to **false** otherwise.

1 **Procedure** EvaluateMove( $m,T$ ):

```
2   /* The evaluation of an infeasible move is set to  $\infty$  */
3    $b \leftarrow \text{true}$ ,  $c \leftarrow -\text{CurrentDuration}(\mathcal{R}_m)$  /*  $\mathcal{R}_m$  denotes the routes impacted by move  $m$  */
4   /* Let  $\mathcal{R}'_m$  denotes the newly created or modified routes obtained after applying move  $m$  */
5   for each route  $r \in \mathcal{R}'_m$  do
6      $t \leftarrow \text{DurationWithoutDetourToCS}(r)$ 
7     if  $t > T$  then return  $(\infty, \text{true})$  /* The move  $m$  is infeasible */;
8      $e^{LB} \leftarrow \text{EnergyWithoutDetourToCS}(r)$ 
9     if  $e^{LB} \leq Q$  then
10       $(c_r, b_r) \leftarrow (t, \text{true})$ 
11    else
12       $t \leftarrow t + \text{MinDurationToDetourToCS}(r)$  /* The computation of the lower bound
13      if  $t > T$  then return  $(\infty, \text{true})$  /* The move  $m$  is infeasible */;
14      if the duration of  $r$  is in the cache memory then  $(c_r, b_r) \leftarrow (\text{DurationFromMemory}(r), \text{true})$  else  $(c_r, b_r) \leftarrow$ 
15       $(t, \text{false})$ 
16    end
17     $c \leftarrow c + c_r$ 
18    if  $b_r = \text{false}$  then  $b \leftarrow \text{false}$ 
19  end
20  if  $c \geq 0$  then  $b \leftarrow \text{true}$  /* The move  $m$  is non-improving */
21  return  $(c, b)$ 
```

---

---

**Algorithm 7:** Checking the violation of capacity constraints for version N of the subproblem

---

**Input** : a list of charging operations  $L$  numbered from 1 to  $n$  ( $L(i)$  denotes the operation at position  $i$  in the list  $L$ ) / an integer number  $C \geq 1$  representing the maximum number of operations that can be scheduled simultaneously  
**Output**: a set containing all the maximal subsets of charging operations leading to a violation of the CS capacity constraint

1 **Procedure** CheckCapacityCut( $O,C$ ):

```
2   Sort the operations in  $L$  in non-decreasing order of their starting time /* if two charging operations have the same
3   starting time, then the charging operation with the minimum duration comes first */
4    $\mathcal{U} \leftarrow \emptyset$  /* subsets of charging operations leading to a violation of the CS capacity constraint*/
5    $\mathcal{U} \leftarrow \{L(1)\}$  /* subset of charging operations currently executed */
6    $k \leftarrow 2$ ,  $\text{excess} \leftarrow \text{false}$ 
7   while  $k \leq n$  do
8     for every operation  $o \in \mathcal{U}$  do
9       if  $\bar{S}(L(k)) \geq \bar{S}(o) + \bar{\Delta}(o)$  then
10        if  $\text{excess}$  then  $\mathcal{U} \leftarrow \mathcal{U} \cup \{o\}$ ,  $\text{excess} \leftarrow \text{false}$ 
11        else  $\mathcal{U} \leftarrow \mathcal{U} \setminus \{o\}$ 
12      end
13    end
14     $\mathcal{U} \leftarrow \mathcal{U} \cup \{L(k)\}$ 
15    if  $|\mathcal{U}| > C$  then  $\text{excess} \leftarrow \text{true}$ 
16  end
17  if  $\text{excess}$  then  $\mathcal{U} \leftarrow \mathcal{U} \cup \{o\}$ 
18  return  $\mathcal{U}$ 
```

---

---

**Algorithm 8:** The fixed part based reasoning algorithm

---

**Input** : a list of charging operations  $O$  numbered from 1 to  $n$  ( $O(k)$  denotes the operation at position  $k$  in the list  $O$ ) /  
an integer number  $C \geq 1$  representing the maximum number of operations that can be scheduled  
simultaneously

**Output:** a set containing all the maximal subsets of charging operations leading to a violation of the CS capacity  
constraint

1 **Procedure** FixedPartsAlgorithm( $O, C$ ):

```
2   Remove from  $O$  all the charging operations such that  $LS(o) \geq ES(o) + \bar{\Delta}(o)$ 
3   Sort the operations in  $O$  in non-decreasing order of starting time /* if two charging operations have the same
   starting time, then the charging operation with the minimum duration comes first */
4    $\mathcal{U} \leftarrow \emptyset$  /* subsets of charging operations leading to a violation of the CS capacity constraint */
5    $U \leftarrow \{O(1)\}$  /* subset of charging operations currently executed */
6    $k \leftarrow 2$ ,  $excess \leftarrow false$ 
7   while  $k \leq n$  do
8     for every operation  $o \in U$  do
9       if  $LS_{O(k)} \geq ES(o) + \bar{\Delta}(o)$  then
10        if  $excess$  then  $\mathcal{U} \leftarrow \mathcal{U} \cup \{U\}$ ,  $excess \leftarrow false$ 
11        else  $U \leftarrow U \setminus \{o\}$ 
12      end
13    end
14     $U \leftarrow U \cup \{O(k)\}$ 
15    if  $|U| > C$  then  $excess \leftarrow true$ 
16  end
17  if  $excess$  then  $\mathcal{U} \leftarrow \mathcal{U} \cup \{U\}$ 
18  return  $\mathcal{U}$ 
```

---

---

**Algorithm 9:** The energy reasoning based algorithm

---

**Input** : a set of charging operations  $O$  / an integer number  $C \geq 1$  representing the maximum number of operations  
that can be scheduled simultaneously

**Output:** a set containing all subsets of charging operations leading to a violation of the CS capacity constraint

1 **Procedure** EnergeticAlgorithm( $O, C$ ):

```
2    $T_1 := \bigcup_{o \in O} (\{ES(o)\} \cup \{ES(o) + \bar{\Delta}(o)\} \cup \{LS(o)\})$ 
3    $T_2 := \bigcup_{o \in O} (\{LS(o) + \bar{\Delta}(o)\} \cup \{ES(o) + \bar{\Delta}(o)\} \cup \{LS(o)\})$ 
4    $T(t) := \bigcup_{o \in O} \{ES(o) + LS(o) + \bar{\Delta}(o) - t\}$ 
5    $T = \{(t_1, t_2) \in T_1 \times T_2\} \cup \{(t_1, t_2) \in T_1 \times T(t_1) : t_1 \in O_1\} \cup \{(t_1, t_2) \in T(t_2) \times T_2 : t_2 \in O_2\}$ 
6    $p_o^+(t_1) = \max\{0, \bar{\Delta}(o) - \max\{0, t_1 - ES(o)\}\}$  /* duration during which  $o$  is executed after time  $t_1$  if it is scheduled
   as soon as possible */
7    $p_o^-(t_2) = \max\{0, \bar{\Delta}(o) - \max\{0, LS(o) + \bar{\Delta}(o) - t_2\}\}$  /* duration during which  $o$  is executed before time  $t_2$  if it is
   scheduled as late as possible */
8    $W(o, t_1, t_2) = \min\{0, t_2 - t_1, p_o^+(t_1), p_o^-(t_2)\}$ 
9   for  $(t_1, t_2) \in T_1 \times T_2$  do
10    if  $t_1 < t_2 \wedge \sum_{o \in O} W(o, t_1, t_2) > C(t_2 - t_1)$  then  $\mathcal{U} \leftarrow \mathcal{U} \cup \{o \in O : W(o, t_1, t_2) > 0\}$ 
11  end
12  return  $\mathcal{U}$ 
```

---



## E Detailed computational results

Based on notation introduced by Montoya et al. (2017), we write each instance using the symbol  $tc\gamma_1c\gamma_2s\gamma_3c\gamma_4\#$  where  $\gamma_1$  is the method used to place the customers (i.e., 0: randomization, 1: mixture of randomization and clustering, 2: clustering),  $\gamma_2$  is the number of customers,  $\gamma_3$  is the number of the CSs,  $\gamma_4$  is 't' if we use a  $p$ -median heuristic to locate the CSs and 'f' otherwise, and  $\#$  is the number of the instance for each combination of parameters (i.e.,  $\# = 0, 1, 2, 3, 4$ ).

### E.1 Detailed results for the E-VRP-NL-C formulations

Tables 10 and 11 report the detailed results obtained by the CS replication-based and path-based formulations on the E-VRP-NL-C. The symbol “-” means that no feasible solution has been found by the solver.

### E.2 Detailed results for the E-VRP-NL

Tables 12 and 13 report the detailed results obtained by our ILS and two methods from the literature on the E-VRP-NL (Montoya et al., 2017; Koç et al., 2018). The best values (objective function value of the BKS and best average value of the objective function obtained over 10 runs) are indicated in boldface.

### E.3 Detailed results for the E-VRP-NL-C

Tables 14, 15, 16, and 17 report the detailed results obtained by our matheuristic on the E-VRP-NL-C instances. The best values (objective function value of the BKS and best average value of the objective function obtained over 10 runs) are indicated in boldface.

Table 10: Detailed results for the CS replication-based and path-based MILP formulations on the 10-customer and 20-customer instances if the number of chargers at each CS is equal to 1.

Instance	$[F^{CSrep}] (\beta = 4)$		$[F^{CSrep}] (\beta = 5)$		$[F^{path}]$	
	Obj	Time	Obj	Time	Obj	Time
tc0c10s2cf1	19.75	112	19.75	123	19.75	9
tc0c10s2ct1	12.30	76	12.30	115	12.30	21
tc0c10s3cf1	19.75	10800	19.75	10800	19.76	832
tc0c10s3ct1	10.80	11	10.80	37	10.80	25
tc1c10s2cf2	9.03	7	9.03	10	9.03	8
tc1c10s2cf3	16.37	49	16.37	77	16.37	46
tc1c10s2cf4	16.10	30	16.10	60	16.10	12
tc1c10s2ct2	10.75	127	10.75	159	10.75	446
tc1c10s2ct3	13.17	41	13.17	54	13.17	6
tc1c10s2ct4	13.83	32	13.83	23	13.83	5
tc1c10s3cf2	9.03	6	9.03	11	9.03	14
tc1c10s3cf3	16.37	1558	16.37	1433	16.37	299
tc1c10s3cf4	14.90	65	14.90	252	14.90	36
tc1c10s3ct2	9.20	243	9.20	956	9.20	272
tc1c10s3ct3	13.02	1105	13.02	515	13.02	156
tc1c10s3ct4	13.21	90	13.21	466	13.21	36
tc2c10s2cf0	21.77	10800	21.83	10800	21.77	988
tc2c10s2ct0	12.45	2391	12.45	3750	12.45	482
tc2c10s3cf0	21.83	10800	21.77	10800	21.77	3444
tc2c10s3ct0	11.51	10800	11.51	10800	11.51	4309
tc0c20s3cf2	-	10800	27.47	10800	27.49	10800
tc0c20s3ct2	17.08	10800	17.45	10800	17.08	10800
tc0c20s4cf2	-	10800	-	10800	27.51	10800
tc0c20s4ct2	17.22	10800	17.35	10800	17.24	10800
tc1c20s3cf1	17.49	10800	17.49	10800	17.49	10800
tc1c20s3cf3	16.80	10800	16.81	10800	16.83	10800
tc1c20s3cf4	17.00	179	17.00	197	17.00	573
tc1c20s3ct1	18.94	10800	18.94	10800	19.39	10800
tc1c20s3ct3	12.60	4315	12.60	1181	12.60	10800
tc1c20s3ct4	16.21	277	16.21	341	16.21	513
tc1c20s4cf1	16.38	10800	16.38	10800	16.38	10800
tc1c20s4cf3	16.80	10800	16.80	10800	16.86	10800
tc1c20s4cf4	17.00	1309	17.00	640	-	10800
tc1c20s4ct1	17.81	10800	17.81	10800	19.12	10800
tc1c20s4ct3	14.43	10800	14.43	10800	14.43	10800
tc1c20s4ct4	17.00	1241	17.00	1184	-	10800
tc2c20s3cf0	25.05	10800	-	10800	24.85	10800
tc2c20s3ct0	25.79	10800	25.80	10800	25.82	10800
tc2c20s4cf0	-	10800	-	10800	24.73	10800
tc2c20s4ct0	-	10800	-	10800	26.36	10800

Table 11: Detailed results for the CS replication-based and path-based MILP formulations on the 10-customer and 20-customer instances if the number of chargers at each CS is equal to 2.

Instance	$[F^{CSrep}] (\beta = 4)$		$[F^{CSrep}] (\beta = 5)$		$[F^{path}]$	
	Obj	Time	Obj	Time	Obj	Time
tc0c10s2cf1	19.75	46	19.75	195	19.75	9
tc0c10s2ct1	12.30	105	12.30	172	12.30	22
tc0c10s3cf1	19.75	10800	19.75	10800	19.75	829
tc0c10s3ct1	10.80	14	10.80	23	10.80	28
tc1c10s2cf2	9.03	6	9.03	2	9.03	9
tc1c10s2cf3	16.37	58	16.37	60	16.37	53
tc1c10s2cf4	16.10	22	16.10	51	16.10	13
tc1c10s2ct2	10.75	139	10.75	166	10.75	535
tc1c10s2ct3	13.17	22	13.17	61	13.17	6
tc1c10s2ct4	13.83	24	13.83	64	13.83	5
tc1c10s3cf2	9.03	2	9.03	25	9.03	16
tc1c10s3cf3	16.37	1328	16.37	2994	16.37	306
tc1c10s3cf4	14.90	59	14.90	235	14.90	43
tc1c10s3ct2	9.20	120	9.20	535	9.20	302
tc1c10s3ct3	13.02	1567	13.02	789	13.02	188
tc1c10s3ct4	13.21	81	13.21	83	13.21	40
tc2c10s2cf0	21.77	10800	21.77	10800	21.77	1224
tc2c10s2ct0	12.45	3305	12.45	3847	12.45	577
tc2c10s3cf0	21.77	10800	21.77	10800	21.77	4002
tc2c10s3ct0	11.51	10800	11.51	10800	11.51	4222
tc0c20s3cf2	-	10800	-	10800	27.49	10800
tc0c20s3ct2	17.08	10800	17.08	10800	17.08	10800
tc0c20s4cf2	-	10800	-	10800	27.51	10800
tc0c20s4ct2	16.99	10800	17.22	10800	17.43	10800
tc1c20s3cf1	17.49	10800	17.49	10800	17.49	10800
tc1c20s3cf3	16.45	10800	16.80	10800	16.83	10800
tc1c20s3cf4	17.00	108	17.00	177	17.00	712
tc1c20s3ct1	18.94	10800	18.94	10800	19.39	10800
tc1c20s3ct3	12.60	529	12.60	4341	12.60	10800
tc1c20s3ct4	16.21	309	16.21	329	16.21	599
tc1c20s4cf1	16.38	10800	16.38	10800	16.38	10800
tc1c20s4cf3	16.80	10800	16.80	10800	17.04	10800
tc1c20s4cf4	17.00	545	17.00	1530	20.06	10800
tc1c20s4ct1	18.02	10800	18.25	10800	19.12	10800
tc1c20s4ct3	14.43	10800	14.43	10800	14.43	10800
tc1c20s4ct4	17.00	1495	17.00	559	-	10800
tc2c20s3cf0	-	10800	24.87	10800	24.85	10800
tc2c20s3ct0	25.79	10800	25.79	10800	25.83	10800
tc2c20s4cf0	24.86	10800	-	10800	24.73	10800
tc2c20s4ct0	-	10800	-	10800	26.36	10800















## References

- Alvarenga, G. B., Mateus, G., and De Tomi, G. (2007). A genetic and set partitioning two-phase approach for the vehicle routing problem with time windows. *Computers & Operations Research*, 34(6):1561–1584.
- Andelmin, J. and Bartolini, E. (2019). A multi-start local search heuristic for the green vehicle routing problem based on a multigraph reformulation. *Computers & Operations Research*, 109:43–63.
- Artigues, C., Michelon, P., and Reusser, S. (2003). Insertion techniques for static and dynamic resource-constrained project scheduling. *European Journal of Operational Research*, 149(2):249–267.
- Baptiste, P., Le Pape, C., and Nuijten, W. (2001). *Constraint-based scheduling: applying constraint programming to scheduling problems*, volume 39. Springer, Boston, MA.
- Bartolini, E. and Andelmin, J. (2017). An exact algorithm for the green vehicle routing problem. *Transportation Science*, 51(4):1288–1303.
- Beek, O., Raa, B., Dullaert, W., and Vigo, D. (2018). An efficient implementation of a static move descriptor-based local search heuristic. *Computers & Operations Research*, 94:1–10.
- Bruglieri, M., Mancini, S., and Pisacane, O. (2019). The green vehicle routing problem with capacitated alternative fuel stations. *Computers & Operations Research*, 112:104759.
- Codato, G. and Fischetti, M. (2006). Combinatorial Benders’ cuts for mixed-integer linear programming. *Operations Research*, 54(4):756–766.
- Desaulniers, G., Errico, F., Irnich, S., and Schneider, M. (2016). Exact algorithms for electric vehicle-routing problems with time windows. *Operations Research*, 64(6):1388–1405.
- Drexl, M. (2012). Synchronization in vehicle routing – a survey of vrps with multiple synchronization constraints. *Transportation Science*, 46(3):297–316.
- El Hachemi, N., Gendreau, M., and Rousseau, L.-M. (2013). A heuristic to solve the synchronized log-truck scheduling problem. *Computers & Operations Research*, 40(3):666–673.
- Erdoğan, S. and Miller-Hooks, E. (2012). A green vehicle routing problem. *Transportation Research Part E: Logistics and Transportation Review*, 48(1):100–114.
- Felipe, A., Ortuño, M. T., Righini, G., and Tirado, G. (2014). A heuristic approach for the green vehicle routing problem with multiple technologies and partial recharges. *Transportation Research Part E: Logistics and Transportation Review*, 71:111–128.
- Froger, A., Mendoza, J. E., Jabali, O., and Laporte, G. (2019). Improved formulations and algorithmic components for the electric vehicle routing problem with nonlinear charging functions. *Computers & Operations Research*, 104:256–294.
- Gnann, T., Funke, S., Jakobsson, N., Plötz, P., Sprei, F., and Bennehag, A. (2018). Fast charging infrastructure for electric vehicles: Today’s situation and future needs. *Transportation Research Part D: Transport and Environment*, 62:314–329.
- Goeke, D. and Schneider, M. (2015). Routing a mixed fleet of electric and conventional vehicles. *European Journal of Operational Research*, 245(1):81–99.
- Grangier, P., Gendreau, M., Lehuédé, F., and Rousseau, L.-M. (2017). A matheuristic based on large neighborhood search for the vehicle routing problem with cross-docking. *Computers & Operations Research*, 84:116–126.
- Grimault, A., Bostel, N., and Lehuédé, F. (2017). An adaptive large neighborhood search for the full truckload pickup and delivery problem with resource synchronization. *Computers & Operations Research*, 88:1–14.

- Hempesch, C. and Irnich, S. (2008). Vehicle routing problems with inter-tour resource constraints. In Golden, B., Raghavan, S., and Wasil, E., editors, *The Vehicle Routing Problem: Latest Advances and New Challenges*, pages 421–444. Springer, Boston, MA.
- Hiermann, G., Hartl, R. F., Puchinger, J., and Vidal, T. (2019). Routing a mix of conventional, plug-in hybrid, and electric vehicles. *European Journal of Operational Research*, 272(1):235–248.
- Hiermann, G., Puchinger, J., Ropke, S., and Hartl, R. F. (2016). The electric fleet size and mix vehicle routing problem with time windows and recharging stations. *European Journal of Operational Research*, 252(3):995–1018.
- Juan, A. A., Mendez, C. A., Faulin, J., de Armas, J., and Grasman, S. E. (2016). Electric vehicles in logistics and transportation: A survey on emerging environmental, strategic, and operational challenges. *Energies*, 9(2):86.
- Keskin, M. and Çatay, B. (2016). Partial recharge strategies for the electric vehicle routing problem with time windows. *Transportation Research Part C: Emerging Technologies*, 65:111–127.
- Keskin, M., Laporte, G., and Çatay, B. (2019). Electric vehicle routing problem with time-dependent waiting times at recharging stations. *Computers & Operations Research*, 107:77–94.
- Koç, Ç., Jabali, O., Mendoza, J. E., and Laporte, G. (2018). The electric vehicle routing problem with shared charging stations. *International Transactions in Operational Research*, 26(4):1211–1243.
- Koç, C. and Karaoglan, I. (2016). The green vehicle routing problem: A heuristic based exact solution approach. *Applied Soft Computing*, 39:154–164.
- Kullman, N., Goodson, J., and Mendoza, J. E. (2018). Dynamic Electric Vehicle Routing: Heuristics and Dual Bounds. Working paper available at <https://hal.archives-ouvertes.fr/hal-01928730>.
- Lam, E. and Van Hentenryck, P. (2016). A branch-and-price-and-check model for the vehicle routing problem with location congestion. *Constraints*, 21(3):1–19.
- Lourenço, H. R., Martin, O. C., and Stützle, T. (2003). Iterated local search. In Glover, F. and Kochenberger, G. A., editors, *Handbook of metaheuristics*, volume 57, pages 320–353. Springer, Boston, MA.
- Montoya, A., Guéret, C., Mendoza, J., and Villegas, J. (2016). A multi-space sampling heuristic for the green vehicle routing problem. *Transportation Research Part C: Emerging Technologies*, 70:113–128.
- Montoya, A., Guéret, C., Mendoza, J. E., and Villegas, J. G. (2017). The electric vehicle routing problem with nonlinear charging function. *Transportation Research Part B: Methodological*, 103:87–110.
- Morais, V. W. C., Mateus, G. R., and Noronha, T. F. (2014). Iterated local search heuristics for the vehicle routing problem with cross-docking. *Expert Systems with Applications*, 41(16):7495–7506.
- Pelletier, S., Jabali, O., and Laporte, G. (2016). 50th anniversary invited article—goods distribution with electric vehicles: Review and research perspectives. *Transportation Science*, 50(1):3–22.
- Pelletier, S., Jabali, O., and Laporte, G. (2018). Charge scheduling for electric freight vehicles. *Transportation Research Part B: Methodological*, 115:246–269.
- Pelletier, S., Jabali, O., Laporte, G., and Veneroni, M. (2017). Battery degradation and behaviour for electric vehicles: Review and numerical analyses of several models. *Transportation Research Part B: Methodological*, 103:158–187.
- Rix, G., Rousseau, L.-M., and Pesant, G. (2015). A column generation algorithm for tactical timber transportation planning. *Journal of the Operational Research Society*, 66(2):278–287.
- Sassi, O. and Oulamara, A. (2014). Joint scheduling and optimal charging of electric vehicles problem. In Murgante, B., Misra, S., Rocha, A. M. A. C., Torre, C., Rocha, J. G., Falcão, M. I., Tanar, D., Apduhan, B. O., and Gervasi, O., editors, *Computational Science and Its Applications – ICCSA 2014*, volume 8580, pages 76–91, Cham. Springer.

- Schiffer, M., Stütz, S., and Walther, G. (2018). Electric commercial vehicles in mid-haul logistics networks. In Pistoia, G. and Liaw, B., editors, *Behaviour of Lithium-Ion Batteries in Electric Vehicles*, pages 153–173. Springer, Cham.
- Schneider, M., Stenger, A., and Goeke, D. (2014). The electric vehicle-routing problem with time windows and recharging stations. *Transportation Science*, 48(4):500–520.
- Subramanian, A., Uchoa, E., and Ochi, L. S. (2013). A hybrid algorithm for a class of vehicle routing problems. *Computers & Operations Research*, 40(10):2519–2531.
- Sweda, T. M., Dolinskaya, I. S., and Klabjan, D. (2017). Adaptive routing and recharging policies for electric vehicles. *Transportation Science*, 51(4):1326–1348.
- Villegas, J. G., Guéret, C., Mendoza, J. E., and Montoya, A. (2018). The technician routing and scheduling problem with conventional and electric vehicle. Working paper available at <https://hal.archives-ouvertes.fr/hal-01813887>.
- Villegas, J. G., Prins, C., Prodhon, C., Medaglia, A. L., and Velasco, N. (2013). A matheuristic for the truck and trailer routing problem. *European Journal of Operational Research*, 230(2):231–244.
- Zachariadis, E. E. and Kiranoudis, C. T. (2010). A strategy for reducing the computational complexity of local search-based methods for the vehicle routing problem. *Computers & Operations Research*, 37(12):2089–2105.