TOWARD EFFICIENT TRANSPORTATION ELECTRIFICATION OF HEAVY-DUTY TRUCKS: JOINT SCHEDULING OF TRUCK ROUTING AND CHARGING

Mikhail A. Bragin, Zuzhao Ye, Nanpeng Yu
Department of Electrical and Computer Engineering
University of California
Riverside, CA, USA
mbragin@engr.ucr.edu, zye066@ucr.edu, nyu@ece.ucr.edu

ABSTRACT

The timely transportation of goods to customers is an essential component of economic activities. However, heavy-duty diesel trucks used for goods delivery significantly contribute to greenhouse gas (GHG) emissions within many large metropolitan areas, including Los Angeles, New York, and San Francisco. To reduce GHG emissions by facilitating freight electrification, this paper proposes Joint Routing and Charging (JRC) scheduling for electric trucks. The objective of the associated optimization problem is to minimize the cost of transportation, charging, and tardiness. A large number of possible combinations of road segments as well as a large number of combinations of charging decisions and charging durations leads to a combinatorial explosion in the possible decisions electric trucks can make. The resulting mixed-integer linear programming (MILP) problem is thus extremely challenging because of the combinatorial complexity even in the deterministic case. Therefore, a Surrogate Level-Based Lagrangian Relaxation (SLBLR) method is employed to decompose the overall problem into significantly less complex truck subproblems. In the coordination aspect, each truck subproblem is solved independently of other subproblems based on the values of Lagrangian multipliers. In addition to serving as a means of guiding and coordinating trucks, multipliers can also serve as a basis for transparent and explanatory decision-making by trucks. Testing results demonstrate that even small instances cannot be solved using the off-the-shelf solver CPLEX after several days of solving. The SLBLR method, on the other hand, can obtain near-optimal solutions within a few minutes for small cases, and within 30 minutes for large ones. Furthermore, it has been demonstrated that as battery capacity increases, the total cost decreases significantly; moreover, as the charging power increases, the number of trucks required decreases as well.

Keywords  Heavy-Duty Vehicles · Transportation Electrification · Surrogate Level-Based Lagrangian Relaxation

1 Introduction

The transportation of goods to customers is an essential economic activity. An effective logistical design is critical in ensuring the delivery of goods on time. However, as a result of the excessive use of transportation in large metropolitan areas, such as Los Angeles, New York, and San Francisco, many of these areas are heavily congested and polluted. In addition to public transportation, commercial transportation, such as large freight trucks that deliver goods, contributes significantly to greenhouse gas emissions. More than 40% of greenhouse gas emissions in California are attributed to the transportation sector, as reported in [1]. Transportation has been widely acknowledged for its environmental impacts (such as climate change due to emissions) [2–6]. Although heavy-duty trucks make up only 1% of all vehicles, their emissions account for 25% of all vehicle emissions [7]. In light of this, electrification of the transportation sector, especially heavy-duty vehicles, offers a promising means of reducing greenhouse gas emissions [4, 8–11] as well as addressing fossil fuel shortages [11].

Despite significant advances in electrifying a range of vehicles, such as passenger cars and transit buses, the availability of electric heavy-duty trucks in the market remains notably limited. There are several reasons for this disparity. First, the energy consumption per mile for heavy-duty trucks is significantly higher than that of passenger cars, requiring...
larger batteries and consequently, driving up initial costs. For example, while a passenger car uses roughly 0.3 kWh per mile, a heavy-duty truck consumes more than 2 kWh [12]. Second, compared to transit buses, heavy-duty trucks are driven considerably farther each year (62,000 miles compared to 43,000 miles, according to [13]). This increased operational range necessitates the fleets of electric heavy-duty trucks to opt for more frequent, higher-powered, or longer-duration charging. Lastly, the charging infrastructure for heavy-duty trucks is largely underdeveloped. The construction of charging stations, requiring high power capacity and substantial space, is likely to be expensive, leading to a limited number of such facilities in the near future. Each station, consequently, is expected to face high demand, thereby leading to potentially long wait times and elevated risks of delivery delays. Given these challenges, it is crucial to develop strategically tailored algorithms for joint routing and charging as a key means for ensuring sustainable operations of heavy-duty trucks upon their electrification.

Existing literature extensively addresses various aspects of electric vehicle (EV) operations. The majority of the research has been conducted on passenger cars and transit buses. The operations of electric passenger cars (or light-duty vehicles) usually fall into the category of electric vehicle routing problems (EVRP) [14,15]. This field focuses on finding optimal routes and charging plans for electric vehicles. Current EVRP research can be generally distinguished by its objective functions and corresponding constraints. The common objective function typically includes operational cost factors such as travel distance [16,17], travel time [18], charging costs [19], energy consumption [20,21], and the mix of two or more factors [19,22,25]. The compulsory constraint is the battery energy level constraint, while common constraints include pickup-and-delivery [17,26] and vehicle loading capacity [16,23,27]. Beyond the common considerations, [28] proposed routing EVs in a power network for vehicle-to-grid (V2G) applications, thus optimizing the timing and location of energy trading. Autonomous EVs, due to their increased flexibility without drivers, hold even more potential for V2G applications [29,30]. Ref. [31] considered the operations of mixed fleets of electric and conventional vehicles, which is particularly meaningful during the transition phase to EVs. Enabling additional customer visits by walking or other means of transportation while the EV is being charged could potentially reduce EV downtime, as suggested by [32]. There are also studies targeting individual EVs instead of the fleet. Ref. [33] studied the routing of individual EVs under an improved Dijkstra algorithm and with forecasting of charging cost. More comprehensive literature reviews can be found in [14,15,34,35]. In a short summary, despite variations in problem settings, the fundamental concerns of the EVRP remain consistent.

The operations of electric transit buses are generally categorized as electric bus scheduling problems (EBSP). In contrast to EVRP, EBSP tends not to emphasize routing since bus routes are typically predetermined. Instead, EBSP concentrates on optimal bus assignment to cover the entire transit system timetable [36,37]. The objective usually encompasses the costs of vehicles and associated charging infrastructure (as the transit system is usually self-contained) [38,39], in addition to the operational costs considered in EVRP. Attention is often given to charging station capacity due to the significant size of transit buses and the substantial costs related to high-power charging [40,41]. There are many interesting works in EBSP. [37] considers the installation of dynamic wireless charging lanes such that the battery size and the station charging time can both be reduced. [42] considers treating electric buses as temporary mobile power sources to provide grid services. Battery degradation is also considered by multiple studies [43,44], as replacing the high-capacity battery of transit buses can be costly.

The operation of electric heavy-duty trucks, while sharing commonalities with passenger cars and transit buses, also exhibits unique characteristics:

- First, the tasks of heavy-duty trucks vary daily and lack fixed timetables, unlike transit buses. This necessitates joint routing and charging scheduling, positioning this problem as an enhanced variant of EVRP rather than EBSP.
- Second, truck operations share a common concern with EBSP regarding charging station capacity due to their substantial dimensions. At any given time, the number of vehicles being charged at a station cannot exceed the number of available chargers. To accommodate this requirement, an additional time dimension is introduced, which is common in EBSP [40,41], but not found with EVRP. Specifically, the time horizon needs to be discretized to enable the modeling of charging station capacity limit in any time period. This additional dimension amplifies the model’s complexity.
- Last but not least, the operation of heavy-duty trucks carries its own unique considerations. These include treating time constraints as tardiness instead of compulsory time windows and modeling the influence of cargo on the energy consumption rate - it necessitates the introduction of additional binary variables.

As a result, the operation of heavy-duty trucks demands a specialized model formulation, distinct from those of transit buses or passenger cars. While the existing literature on EVRP has also considered certain unique features of heavy-duty trucks ( [24] considers charging station capacity and [45] considers the impacts of cargo on the energy consumption rate), to the best of our knowledge, no comprehensive study has been found that encompasses all these aspects. Past studies specifically targeting heavy-duty trucks have exhibited limitations. [46] sought to minimize elec-
tric freight truck charging costs but assumed fixed routes, optimizing only the charging schedule without considering route alterations. [47] employed a bi-level heuristic for joint charging and routing of electric trucks but overlooked the essential factor of charging station capacity. In summary, there is a pressing need for a comprehensive and efficient framework for the joint charging and routing of electric heavy-duty trucks.

Regarding solution methods, EBSPs are generally solved exactly by commercial solvers [37], as they naturally skip the complexity added by routing on road networks. EVRP, on the other hand, can only be solved exactly for small instances due to their additional routing requirements. As such, EVRP generally relies on heuristic or meta-heuristic methods for larger instances [14]. Given the wide adoption of meta-heuristic methods, we will provide a brief overview. Commonly used meta-heuristic methods include genetic algorithms [18, 48, 49], variable neighborhood search [23, 27, 45, 50], ant colony optimization [16, 41], iterated local search [20, 25, 32], tabu search [17, 19, 52], simulated annealing [24, 26, 53], large neighborhood search [22, 54, 55], and differential evolution algorithms [21]. While the above methods represent a broad range of solution techniques, the list is by no means exhaustive. A common element among these methods is the use of heuristics to reduce the solution time of EVRP which features high combinatorial complexity. However, due to their heuristic nature, these algorithms cannot be systematically improved based on theory. Moreover, these meta-heuristics largely rely on case-by-case developed, hand-crafted heuristics, meaning they are not always adaptable to other variants. In contrast, although Lagrangian Relaxation methods may also require heuristics, the solutions obtained can be continuously improved. This improvement is achieved through efficient coordination using Lagrangian multipliers. Lagrangian Relaxation was indeed considered by [29, 52, 56] to reduce the complexity of EVRP. However, their final solution method did not consider the coordination of Lagrangian multipliers; instead, they reverted to the use of traditional meta-heuristics. As such, the benefit of Lagrangian Relaxation is not fully utilized.

We present an overview of the typical characteristics of the operational problems for various types of vehicles in Table 1, based on our literature review. Note that Table 1 does not aim to provide an exhaustive summary. Instead, it seeks to highlight common features across similar types of vehicles.

<table>
<thead>
<tr>
<th>Vehicle type</th>
<th>Objective function</th>
<th>Routing</th>
<th>Charging scheduling</th>
<th>Fleet or individual</th>
<th>Constraints</th>
<th>Solution method</th>
<th>Heuristic</th>
<th>Meta-heuristic</th>
<th>Exact</th>
<th>Meta-heuristic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passenger car</td>
<td>Op</td>
<td>✓</td>
<td>✓</td>
<td>Both</td>
<td>Time(2)</td>
<td>Heuristic</td>
<td>Exact</td>
<td>Meta-heuristic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transit bus</td>
<td>Op + Veh + Infra</td>
<td>✓</td>
<td>✓</td>
<td>Fleet</td>
<td>Energy</td>
<td>Meta-heuristic</td>
<td>Exact</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heavy-duty truck</td>
<td>Op</td>
<td>✓</td>
<td>✓</td>
<td>Fleet</td>
<td>Pickup-delivery</td>
<td>Meta-heuristic</td>
<td>Exact</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Station capacity</td>
<td></td>
<td>Exact</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cargo(3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(1) Op: Operational costs, e.g. travel distance/time and charging costs; Veh: Vehicle procurement cost; Infra: Charging infrastructure cost.
(2) TW: Time window; TB: Trip-based constraint, i.e. the arrival time of a trip should be ahead of its connecting trip; TD: Tardiness or soft time window, appearing as a penalty term.
(3) The loading/unloading of cargo will create two levels of energy consumption rates that require additional binary constraints to handle.

The rest of the paper is organized in the following manner. We present a novel formulation of a Joint heavy-duty vehicle fleet Routing and Charging (JRC) problem in Section 2. Compared with EBSP, this formulation introduces additional routing requirements; Compared with general EVRP, this formulation adds a completely new time dimension to decision variables to enable the modeling of charging station capacity limit, similar to that of EBSP. The loading/unloading of cargo and their impacts on the energy consumption rate are also modeled through a set of additional binary variables. We consider the new formulation as an enhanced variant of EVRP.

To effectively manage the increased complexity introduced by the innovative problem formulation and to address the limitations of existing methodologies, we utilize the principles behind the Surrogate “Level-Based” Lagrangian Relaxation (SLBLR) approach [57], the nuances of which are elaborated in Section 3. The SLBLR possesses several key desirable properties:

---

Page 3 of 27
1. It drastically reduces the combinatorial complexity via decomposition into truck subproblems; the method essentially “reverses” the inherent combinatorial complexity of the underlying problem;

2. The surrogate component of the methodology allows for bypassing the need for achieving optimal solutions for the truck subproblems at each iteration as long as the “surrogate optimality condition” is satisfied. This feature, which is lacking in the standard Lagrangian Relaxation appears to be the dominant as will be demonstrated in Section 4.

3. The “level-based” coordination nature of our method exploits the linear convergence rate (fastest possible in the dual space). While other methods [58] also utilize the “level-based” ideas, their level adjustment is heuristic in nature, while our adjustment is informed and decision-based leading to a faster adjustment of level values thus allowing us to exploit linear convergence potential more efficiently;

4. Penalty terms are used for constraint violations to accelerate convergence and are even capable of “pushing” the total constraint violation to zero thereby avoiding the need for coding other heuristics for primal solution recovery.

Besides the above properties, the SLBLR method inherently benefits from the following features:

1. Flexibility and Adaptability. The SLBLR method exhibits model-agnostic capabilities. In essence, the same algorithm can be easily adapted to different model formulations, ensuring its versatility and adaptability across a wide array of problem scenarios;

2. Continuous Improvement. The SLBLR approach enables a continuous refinement of solutions. Unlike meta-heuristic methods employed for EVRP, the SLBLR technique facilitates continual improvements of solution achieved through the efficient coordination of Lagrangian multipliers;

3. Economic Viability and Explainability. The method’s price-based feature stems from the economic theory, wherein Lagrangian multipliers are interpreted as “shadow prices” that adjust according to the “supply and demand” principle, ensuring that the obtained solutions are economically viable. In terms of the JRC problem, the “demand” is the number of trucks needing to charge at a particular node of the road network, and the “supply” is the number of chargers. When the “demand” exceeds the “supply” (charging stations are oversubscribed), “prices” (not to be confused with the charging price) increase thereby discouraging trucks from making less “economically viable” decisions, e.g., from driving to an oversubscribed charging station thereby losing time and facing a higher penalty for late shipments. Multipliers thus possess the intuitive explainability feature behind the underlying decision making such as the grounds for the re-routing of trucks along the longer paths, which lead to lower overall costs. The same logic can be applied to the actual demand and supply of goods to be transported;

4. Smooth Convergence. The SLBLR method (like other LR-based methods) is an iterative approach and optimal multipliers are obtained by taking a series of steps along multiplier-updating directions. While traditionally the subgradient directions were used, surrogate subgradients were shown to be more beneficial for saving computational effort and leading to smoother convergence. The choice of step sizes is not only crucial for guaranteeing convergence but also important for achieving a high rate of convergence, e.g., fast iteration-wise convergence. The level-based feature of the method rests upon the fundamental principles Polyak stepsize [59]; moreover, our method resolves the issue of the unavailability of knowledge about the optimal value required for fast convergence within [59].

Recently, although within a different problem context, SLBLR [57] demonstrated fast convergence, and several discrete programming problems have been solved while beating the default CPLEX by several orders of magnitude in terms of CPU time.

A series of case studies in Section 4 demonstrate that fast coordination of truck subproblems through the SLBLR method yields near-optimal feasible solutions for problems with different sizes: from five trucks to fifty. In addition, a financial impact analysis is conducted, demonstrating that increased capacity of truck batteries leads to a significant reduction in costs; meanwhile, increased charger power also results in fewer trucks being required to transport goods.

2 Problem Description and Formulation

Description. Consider a transportation network with

1. A set of nodes \( n \in N \);
2. A set of road segments \( r \in R \) with each segment characterized by a “starting node” \( s(r) \) and “ending node” \( e(r) \);
3. A set of heavy-duty electric vehicles (trucks) \( v \in V \);

\[ \text{1Here convergence can be understood both in dual and, perhaps, more importantly in the primal space, since the penalties allow to obtain feasible solutions.} \]

\[ \text{2Note that starting and ending nodes are not unique for a road segment since the road segments are generally bi-directional.} \]
1. Each truck \( v \) is designated to a certain depot \( n^\text{depot}_v \) and a certain port \( n_{v}^\text{port} \). Whether cargo needs to be delivered from a depot to a port or vice versa is determined through optimization as explained above; 

2. Each truck is expected to be fully charged overnight, i.e., the initial charge is 100%. This assumption is not strict, since the methodology developed further will be able to handle any level of the initial charge, if feasible; 

3. Each port is a node \( n^\text{port} \) whereby cargo is loaded, and each port is equipped with a charging station; 

4. Each depot is a node \( n^\text{depot} \) whereby the cargo is unloaded, and each depot is also equipped with a charging station; 

5. Other nodes of the network \( n \in N \setminus \{N^\text{port} \cup N^\text{depot} \} \) may or may not contain a charging station; 

6. Driving along a road segment \( r \) is non-preemptive, that is, once started driving at time \( p \), the truck will arrive at the end of the road segment after \( T^\text{trvl}_{r,p} \) time periods; 

7. Each truck \( v \) can take several trips (not necessarily \( T \)) during a scheduling horizon as long as the last trip ends at \( n^\text{depot}_v \); 

8. The increase in the truck battery’s state of charge is a linear function of charging time; 

Figure 1 schematically illustrates one possible routing-charging scenario of one truck (truck 1). After overnight charging at the depot (charging station located at node \( n^\text{depot}_1 = 1 \)), the truck sets out to pick up the cargo at the port. Having a sufficient amount of charge, the truck drives past the charging station at node \( n^\text{chrg}_1 = 2 \). At node \( n^\text{port}_1 = 3 \), the truck loads (here the loading, as well as unloading times, are assumed negligible). After loading, the truck chooses road segment \( 3, 4 \). At node \( n^\text{chrg}_4 = 4 \), the truck needs to decide whether to charge at \( n^\text{chrg}_5 = 2 \) or \( n^\text{chrg}_5 = 5 \). In this example, it is assumed that the charging station \( n^\text{chrg}_5 = 5 \) is oversubscribed leading to a potentially long waiting time. The truck then travels to node \( n^\text{chrg}_2 = 2 \) to charge. After charging, the truck returns to the depot.

Figure 1: A potential routing and charging scenario for a sample electric truck.

The difference between \( n^\text{depot}_v \) and \( n^\text{depot}_w \) is that \( n^\text{depot}_v \) is a dummy index from a set \( N^\text{depot} \) whereas \( n^\text{depot}_w \) is a specific depot node where truck \( v \) is located.
The goal is to transport all goods from the depot nodes to the port nodes (and vice versa) while co-minimizing the total charging and labor costs, as well as the tardiness penalties due to goods shipped late. To achieve the above goal, the truck needs to decide when and which road segment to take, which charging station to choose, and for how long. The cargo picked up at the port needs to be unloaded at the depot. The above process needs to be repeated until the demand at the depot is satisfied. In the same fashion, the demand at the port is satisfied. These considerations will be mathematically formulated next.

The rest of the section is organized as follows. In subsection 2.1 constraints for scheduling the movement as well as the charging of a heavy-duty truck within one-way trips are modeled. Loading/unloading constraints to connect one-way trips are introduced in subsection 2.2. The multiple-truck scenario will then be considered in subsection 2.3 by introducing charging station “capacity” constraints.

### 2.1 Trip-Wise Constraints

To capture trucks traveling and charging within a one-way trip, several sets of binary decision variables are introduced:

1. To capture the beginning of a trip, let $x_{v,n,p,t}^{trvl,dprt} = 1$ if truck $v$ chooses to depart from node $n$ at time $p$ at trip $t$ and $x_{v,n,p,t}^{trvl,dprt} = 0$ otherwise;
2. The “arrival” binary decision variable $x_{v,n,p,t}^{trvl,arvl}$ is similarly defined;
3. To capture the status of charging, let $x_{v,n,chrg,p,t}^{chrg} = 1$ if truck $v$ chooses to charge at node $n$ during time period $p$ and $x_{v,n,chrg,p,t}^{chrg} = 0$ otherwise;
4. To capture beginning and completion of charging, let $x_{v,n,chrg,bgn,p,t}^{chrg,bgn} = 1$ and $x_{v,n,chrg,cmplt,p,t}^{chrg,cmplt} = 1$ if truck $v$ began charging during trip $t$ at node $n$ at time $p$ and ended charging at time $p$, respectively.

**Truck Availability.** For every truck $v$, the first trip $t = 1$ starts at a depot $n = n_{v,depot}$ after the time $a_{v,depot}$. To capture this condition (as well as several conditions thereafter), binary variable $x_{v,t}^{trip}$ is introduced:

$$
x_{v,1}^{trip} = 1 \Rightarrow d_{v,n_{v,depot},1} \geq a_{v,depot}, \forall (v \in \mathcal{V}),
$$

where $d_{v,n,depot,1}$ is the departure time.

**Departure-Arrival Relationship.** To make sure that truck $v$ can only depart from at most one node during every trip $t$, the following set of constraints is introduced:

$$
\sum_{n \in \mathcal{N}} x_{v,n,p,t}^{trvl,dprt} \leq 1, \forall (v \in \mathcal{V}, p \in \mathcal{P}, t \in \mathcal{T}).
$$

Moreover, truck $v$ can only depart once, therefore, summation over the time periods is also required:

$$
\sum_{p \in \mathcal{P}} x_{v,n,p,t}^{trvl,dprt} \leq 1, \forall (v \in \mathcal{V}, n \in \mathcal{N}, t \in \mathcal{T}).
$$

The corresponding departure time $d_{v,n,t}$ is determined through the following set of constraints:

$$
\sum_{p \in \mathcal{P}} p \cdot x_{v,n,p,t}^{trvl,dprt} = d_{v,n,t}, \forall (v \in \mathcal{V}, n \in \mathcal{N}, t \in \mathcal{T}).
$$

The above constraints (2)-(4) also hold for the binary arrival indicator $x_{v,n,p,t}^{trvl,arvl}$ and integer arrival time $a_{v,n,t}$ decision variables and for brevity are not shown.

Once truck $v$ departs from node $s(r)$ (starting node of road segment $r$), it needs to arrive at one of the nodes $e(r)$ (ending node of road segment $r$) $T_{r,p}^{trvl}$ units of time later (due to non-preemptiveness), which is captured through the following set of constraints:

$$
x_{v,n,p,t}^{trvl,dprt} = 1 \Rightarrow \sum_{r \in \mathcal{R}: s(r) = n} x_{v,e(r),p+T_{r,p}^{trvl} - 1,t}^{trvl,arvl} = 1, \forall (v \in \mathcal{V}, n \in \mathcal{N}, t \in \mathcal{T}, p \in \mathcal{P}).
$$

The above relationship goes both ways: if truck $v$ arrived at node $n$, it must have departed in the past $T_{r,p}^{trvl}$ units of time ago from one of the nodes directly connected by road segments with node $n$:

$$
x_{v,n,p,t}^{trvl,arvl} = 1 \Rightarrow \sum_{r \in \mathcal{R}: e(r) = n} x_{v,s(r),p-T_{r,p}^{trvl} + 1,t}^{trvl,dprt} = 1, \forall (v \in \mathcal{V}, n \in \mathcal{N}, t \in \mathcal{T}, p \in \mathcal{P}).
$$
In the above constraints, one unit of time is added or subtracted from the travel time as appropriate because of the discrete nature of the time periods: departure is understood at the top of the time period while arrival is understood at the end of the time period. For example, if the truck departed at the beginning of time 3 and traveled 3 units of time, then the arrival will be at 3 + 3 - 1 = 5, that is, the truck travels during three time periods 3, 4, and 5.

Moreover, once truck $v$ arrives at node $n$, it needs to depart from $n$ in the future unless truck $v$ reaches the destination in (7). The following sets of constraints, capture this relationship:

$$x_{v,n,p,t}^{\text{trvl,arvl}} = 1 \Rightarrow \sum_{p' \in P, p' \geq p+1} x_{v,n,p,t}^{\text{trvl,drptr}} = 1, \forall (v \in V, n \in N \setminus \{ n_v^{\text{port}} \cup n_v^{\text{depot}} \}, t \in T, p \in P).$$

### Charging and Discharging

While travelling, truck $v$ discharges at rate $\Delta s_{v,e}^{\text{dchrg}}$ per time period, if loaded, and with rate $\Delta s_{v,e}^{\text{dchrg,empty}}$%, if empty. To capture the effect of the above discharge rates, binary decision variables $x_{v,t}^{ld}$ are introduced to indicate whether truck $v$ is loaded during trip $t$ ($x_{v,t}^{ld} = 1$) or empty ($x_{v,t}^{ld} = 0$). With the help of the above-defined “departure” and “arrival” decision variables, the state of charge becomes:

$$x_{v,n,p,t}^{\text{trvl,drptr}} = 1 \land x_{v,e(r),p+T_{r,p}^{\text{trvl}}-1,t}^{\text{trvl,arvl}} = 1 \land x_{v,t}^{ld} = 1 \Rightarrow$$

$$s_{v,e(r),t} = s_{v,n,t}^{\text{chrg}} - \frac{\Delta s_{v,e}^{\text{dchrg,ldd}}}{100} \cdot T_{r,p}^{\text{trvl}} \cdot x_{v,e(r),p+T_{r,p}^{\text{trvl}}-1,t}^{\text{trvl,arvl}}$$

$$\forall (v \in V, n \in N, r \in R : s(r) = n, t \in T, p \in P : p + T_{r,p}^{\text{trvl}} - 1 \leq P),$$

where $s_{v,n,t}^{\text{chrg}} \in [0, 1]$ is a continuous decision variable capturing the state of charge for the battery of truck $v$ at node $n$ during trip $t$; $\land$ denotes the conjunction (also referred to as the “logical AND”). Correspondingly, if truck $v$ is empty during trip $t$ then the above constraint becomes:

$$x_{v,n,p,t}^{\text{trvl,drptr}} = 1 \land x_{v,e(r),p+T_{r,p}^{\text{trvl}}-1,t}^{\text{trvl,arvl}} = 1 \land x_{v,t}^{ld} = 0 \Rightarrow$$

$$s_{v,e(r),t} = s_{v,n,t}^{\text{chrg}} - \frac{\Delta s_{v,e}^{\text{dchrg,empty}}}{100} \cdot T_{r,p}^{\text{trvl}} \cdot x_{v,e(r),p+T_{r,p}^{\text{trvl}}-1,t}^{\text{trvl,arvl}}$$

$$\forall (v \in V, n \in N, r \in R : s(r) = n, t \in T, p \in P : p + T_{r,p}^{\text{trvl}} - 1 \leq P).$$

If truck $v$ arrives at node $n^{\text{chrg}}$ equipped with chargers, it can make a decision to charge; such a decision is captured through a binary variable $x_{v,n^{\text{chrg}},p,t}^{\text{chrg}}$. Assuming that $\Delta s_{v,e}^{\text{chrg}}$% is the charge rate per time period, (8) is modified as:

$$x_{v,n,p,t}^{\text{trvl,drptr}} = 1 \land x_{v,e(r),p+T_{r,p}^{\text{trvl}}-1,t}^{\text{trvl,arvl}} = 1 \land x_{v,t}^{ld} = 1 \Rightarrow$$

$$s_{v,e(r),t} = s_{v,n,t}^{\text{chrg}} - \frac{\Delta s_{v,e}^{\text{dchrg,ldd}}}{100} \cdot T_{r,p}^{\text{trvl}} \cdot x_{v,e(r),p+T_{r,p}^{\text{trvl}}-1,t}^{\text{trvl,arvl}} + \sum_{p' \in P} \frac{\Delta s_{v,e}^{\text{chrg}}}{100} \cdot x_{v,e(r),p',t}^{\text{chrg}}$$

$$\forall (v \in V, n \in N, r \in R : s(r) = n, t \in T, p \in P : p + T_{r,p}^{\text{trvl}} - 1 \leq P).$$

In the above equation $e(r)$ needs to be a member of a set $N^{\text{chrg}}$. Equation (9) is modified in the same way.

In (10), the amount of energy charged is determined through the number of time periods used for charging. However, the appropriate timing for charging needs to be determined to make sure that electric truck $v$ can charge only if it arrived at and departed from node $n^{\text{chrg}}$. Therefore, the following constraints are introduced:

$$\sum_{p=1}^{p-1} x_{v,n^{\text{chrg}},p',t}^{\text{trvl,arvl}} \geq x_{v,n^{\text{chrg}},p,t}^{\text{chrg}}, \forall (v \in V, n^{\text{chrg}} \in N^{\text{chrg}}, t \in T, p \in P).$$

In the equation above, the upper limit of summation $p - 1$ is once again used because of the discrete nature of the time periods; if truck $v$ arrives at the end of the time period $p - 1$, truck $v$ can start charging at time $p$.

The charging will not be possible after the departure, which is captured through the following set of constraints:

$$1 - \sum_{p'=1}^{p} x_{v,n^{\text{chrg}},p',t}^{\text{trvl,drptr}} \geq x_{v,n^{\text{chrg}},p,t}^{\text{chrg}}, \forall (v \in V, n^{\text{chrg}} \in N^{\text{chrg}}, t \in T, p \in P).$$

Note that after arriving at the destination, the truck may still depart. However, the departure will happen during the next trip $t + 1$, which will be discussed in subsection 2.3.
If truck $v$ departs at time $p$, it is naturally no longer eligible for charging starting at time $p$.

The beginning $b_{v,n,chrg,p,t}$ and completion $c_{v,n,chrg,p,t}$ times of charging are captured by introducing binary variables $x_{v,n,chrg,bgn,p,t}$ and $x_{v,n,chrg,cmplt,p,t}$ in the same way as departure times are captured within (3). The same relations as in (2) and (3) hold for $x_{v,n,chrg,bgn,p,t}$ and $x_{v,n,chrg,cmplt,p,t}$ as well. These binary variables are linked to $x_{v,n,chrg,p,t}$ in the following ways:

$$x_{v,n,chrg,bgn,p,t} \geq x_{v,n,chrg,bgn,p,t} = 1, \forall (v \in V, n \in N_{chrg}, t \in T, p \in P \setminus \{1\}),$$

$$x_{v,n,chrg,cmplt,p,t} \geq x_{v,n,chrg,cmplt,p,t} = 1, \forall (v \in V, n \in N_{chrg}, t \in T, p \in P \setminus \{1\}).$$

**Unloading Start Time Constraints.** Since the goal is to co-optimize charging cost and the tardiness, as stated at the beginning of the section, unloading start times will be captured through the use of the integer variables $u_{v,t,n,pr}$:

$$x_{v,t} = 1 \Rightarrow a_{v,t,n,pr} = a_{v,t,n}, \forall (v \in V, n \in \{N_{port} \cup N_{depot}\}, t \in T, pr \in PR).$$

If truck $v$ is loaded at a destination – at the end of trip $t$, and the logical constraint (15) determines the unloading start time at arrival.

### 2.2 Trip-Connecting Constraints

The number of trips that need to be taken is decided through optimization. To capture whether a particular trip is taken, binary variables are introduced $x_{v,t}$, if truck $v$ takes trip $t$ and $x_{v,t} = 0$, otherwise. If truck $v$ is to take the next trip $t + 1$ (which is decided either at a depot $n_{depot}$ or at a port $n_{port}$), then the beginning time of the next trip ($t + 1$) will relate to the arrival time as well as the charging completion time of the previous trip ($t$) in the following ways:

$$x_{v,t+1} = 1 \land \sum_{p \in P} x_{v,n,p,t} = 0 \Rightarrow d_{v,n,t+1} \geq a_{v,n,t} + 1, \forall (v \in V, n \in \{n_{depot}, n_{port}\}, t \in T \setminus \{T\}),$$

$$x_{v,t+1} = 1 \land \sum_{p \in P} x_{v,n,p,t} = 1 \Rightarrow d_{v,n,t+1} \geq c_{v,n,t} + 1, \forall (v \in V, n \in \{n_{depot}, n_{port}\}, t \in T \setminus \{T\}).$$

The above constraints ensure that if truck $v$ is not charging (per (16)), then the next trip $t + 1$ may start immediately after the completion of the previous trip $t$; if the truck needs to charge (per (17)), then the next trip $t + 1$ may start immediately after the completion of the charging within previous trip $t$.

Moreover, at a port ($n = n_{port}$), truck $v$ has to depart to return to a depot, regardless of whether the truck is loaded or not, which is captured through the following set of constraints:

$$x_{v,t} = 1 \land \sum_{p \in P} x_{v,n,p,t} = 0 \Rightarrow d_{v,n,t+1} \geq a_{v,n,t} + 1, \forall (v \in V, n = n_{port}, t \in T),$$

$$x_{v,t} = 1 \land \sum_{p \in P} x_{v,n,p,t} = 1 \Rightarrow d_{v,n,t+1} \geq c_{v,n,t} + 1, \forall (v \in V, n = n_{port}, t \in T).$$

Since ports and depots are equipped with chargers, truck $v$ can make a decision to charge. In the next trip, the charging levels are determined as:

$$x_{v,t+1} = 1 \Rightarrow s_{v,n,t+1} = s_{v,n,t}, \forall (v \in V, n \in \{n_{depot}, n_{port}\}, t \in T \setminus \{T\}),$$

$$x_{v,t} = 1 \Rightarrow s_{v,n,t+1} = s_{v,n,t}, \forall (v \in V, n \in n_{port}, t \in T).$$

Note that charging at nodes $\{n_{depot}, n_{port}\}$ is already accounted for during trip $t$ within (8)–(10).

To ensure that a truck $v$ is not loaded without being scheduled for a trip, we utilize the binary variables $x_{v,t}$, previously introduced in equations (8)–(10). The following constraint set is introduced to maintain this logical consistency:

$$x_{v,t} \leq x_{v,t} = 1 \forall (v \in V, t \in T, pr \in PR).$$

Besides (8)–(10) and the above (22), $x_{v,t}$ will be used to capture the demand as explained ahead in subsection 2.3.

Lastly, to ensure the contiguity of trips, the following set of constraints is introduced:

$$x_{v,t+1} \leq x_{v,t} = 1 \forall (v \in V, t \in T \setminus \{T\}).$$

---

5. Here and throughout the rest of the paper, the “unloading start times” is used to indicate when unloading starts rather how long unloading lasts. As assumed earlier, the unloading duration is negligible.
2.3 Truck-Coupling Constraints

Demand Constraints. Loaded trucks unload cargo at the end of the trip. For inbound trips (to the depot), the unloaded cargo \( p_{n_{\text{depot}}} \) will go towards satisfying the demand \( D_{n_{\text{depot}}, p_{n_{\text{depot}}}} \) at the depot \( n_{\text{depot}} \); for outbound trips (out from the depot), the unloaded cargo \( p_{n_{\text{port}}} \) will go towards satisfying the demand \( D_{n_{\text{port}}, p_{n_{\text{port}}}} \) at the port \( n_{\text{port}} \):

\[
\sum_{t \in T^{\text{inbound}} \cap T, v \in \Omega_{p_{n_{\text{port}}}, n_{\text{depot}}}} x_{v,t}^{ld} = D_{n_{\text{port}}}, \forall (n_{\text{port}} \in N^{\text{port}}, p_{n_{\text{port}}} \in PR),
\]

\[
\sum_{t \in T^{\text{outbound}} \cap T, v \in \Omega_{p_{n_{\text{depot}}}, n_{\text{port}}}} x_{v,t}^{ld} = D_{n_{\text{depot}}}, \forall (n_{\text{depot}} \in N^{\text{depot}}, p_{n_{\text{depot}}} \in PR).
\]

Here \( T^{\text{inbound}} \) is a subset of inbound trips (e.g., trips with odd numbers), and \( T^{\text{outbound}} \) is a subset of outbound trips. Within (24), the summation is over a set of trucks \( \Omega_{p_{n_{\text{port}}}, n_{\text{depot}}} \) eligible to deliver product \( p_{n_{\text{port}}} \) to port \( n_{\text{port}} \) toward satisfaction of demand \( D_{n_{\text{depot}}} \) if truck \( v \) is loaded \( (x_{v,t}^{ld} = 1) \). Within (25), the summation is over a set of trucks \( \Omega_{p_{n_{\text{depot}}}, n_{\text{port}}} \) eligible to deliver product \( p_{n_{\text{depot}}} \) from port \( n_{\text{port}} \) to depot \( n_{\text{depot}} \) toward satisfaction of demand \( D_{n_{\text{depot}}} \) if truck \( v \) is loaded \( (x_{v,t}^{ld} = 1) \).

Charging Station Capacity Constraints. To avoid the situation that more trucks are charging at the same time \( p \) at node \( n_{\text{chrg}} \) than the number of chargers \( C_{n_{\text{chrg}}} \), the following "charging station capacity" constraint is introduced:

\[
\sum_{v \in V, t \in T} x_{v,t}^{chrg} \leq C_{n_{\text{chrg}}}, \forall (n_{\text{chrg}} \in N^{\text{chrg}}, p \in PR).
\]

In (26), the summation is over trips and trucks since a given point in time \( p \) may correspond to different trips for different trucks.

Product Unloading Constraints and Tardiness. The unloading start times \( u_{v,t, p_{r_{p}}} \) that are captured in (15), are used to determine the latest unloading time and tardiness. The latest unloading time is determined as:

\[
\pi_{p_{r}} \geq u_{v,t, p_{r}}, \forall (v \in V, t \in T, p_{r} \in PR).
\]

The tardiness is then defined as:

\[
\pi_{p_{r}} - \text{due}_{p_{r}} \leq \text{tard}_{p_{r}}, \text{tard}_{p_{r}} \geq 0, \forall (p_{r} \in PR).
\]

Tardiness is only positive if the unloading start time is later than the due time, otherwise, the tardiness should be zero. Both situations are materialized through optimization, which is presented ahead.

2.4 Objective Function.

The objective function is to minimize the total labor, charging, and tardiness costs as follows:

\[
\min \left\{ \sum_{v \in V} O_{v}(d, \pi, x) + \sum_{p_{r} \in PR} O_{p_{r}}(\text{tard}) \right\} = \min \left\{ \sum_{v \in V} C_{\text{labr}} \cdot \left( \pi_{v,n_{\text{depot}}} - d_{v,n_{\text{depot}}, 1} \right) + \right. \]

\[
\left. \sum_{v \in V, n_{\text{chrg}} \in N^{\text{chrg}}, t \in T, p_{r} \in PR} C_{\text{chrg}}^{p_{r}} \cdot x_{v,n_{\text{chrg}}, p_{r}}, t + \sum_{p_{r} \in PR} C_{\text{tard}}^{p_{r}} \cdot \text{tard}_{p_{r}} \right\}.
\]

Here \( \pi_{v,n_{\text{depot}}} \) is introduced to capture the latest arrival time of vehicle \( v \) at the depot \( n_{\text{depot}} \) as:

\[
\pi_{v,n_{\text{depot}}} \geq a_{v,n_{\text{depot}}, 1}, \forall (t \in T).
\]

Accordingly, the term \( \left( \pi_{v,n_{\text{depot}}} - d_{v,n_{\text{depot}}, 1} \right) \) is the total time a driver spends on the road (difference between the completion of the last trip taken and the departure during the first trip). The total labor and charging costs are collectively represented as an additive form \( \sum_{v \in V} O_{v}(d, \pi, x) \) in terms of trucks \( v \). The drastic reduction of complexity through the formulation of much smaller and much less complex truck subproblems will be exploited in Section 3.

To avoid the unnecessary increase in complexity, the maximum number of trips as well as the number of time periods need to be appropriately selected as discussed next.

\footnote{Here \( n \) must be equal to \( n \) by assumption 1 stated in the beginning of the Section.}
2.5 Selection of the Maximum Number of Trips and Time Periods

The number of \( T \) is generated based on demand and supply. For instance, consider a scenario where the demand at a port is 18, the demand at a warehouse is 12, and there are 10 trucks. To avoid the unnecessary increase in complexity, the number of round trips is set to 2, accordingly, the number of individual trips is 4. One possible solution is that 10 trucks pick up the cargo demanded during their first trip, and 8 trucks need a second trip to pick up the remaining 8 units of cargo. Therefore, we use a general formula: \( 2 \cdot \max\{P_{\text{prndepot}}, D_{\text{pr}, \text{nport}}\} \) to determine the maximum number of trips. The quantity under the “ceiling” operator is the number of round trips, and the factor of 2 arises to compute the total number of trips.

The determination of the number of time periods follows two main guidelines: to avoid too few time periods, which may result in infeasibility, and too many time periods, which may exacerbate combinatorial complexity. We heuristically estimate the longest path without loops and driving back from a port to a depot. The duration of this path is then multiplied by the largest number of trips defined above.

3 Solution Methodology

This section presents a novel use of the Surrogate “Level-Based” Lagrangian Relaxation (SLBLR) approach [57]. This method, enhanced by penalties, is employed to optimize the nonconvex dual function as well as to recover primal feasibility, aiming to solve the JRC problem formulated in the previous section [2]. Our approach was selected to reduce combinatorial optimization problems, and 2) employ a “surrogate” component that avoids the need for optimal subproblem solutions at each iteration for truck subproblems. This surrogate feature significantly decreases computational efficiency and to efficiently coordinate truck subproblems while facilitating faster convergence by employing “Level-Based” functionality.

The algorithm’s key strengths lie in its ability to 1) “reverse” combinatorial complexity - a desirable feature for complex optimization problems, and 2) employ a “surrogate” component that avoids the need for optimal subproblem solutions at each iteration for truck subproblems. This surrogate feature significantly decreases computational effort and alleviates multiplier zigzagging, offering an advantage over conventional LR methods. Moreover, traditional solutions at each iteration for truck subproblems. This surrogate feature significantly decreases computational efficiency and to efficiently coordinate truck subproblems while facilitating faster convergence by employing “Level-Based” functionality.

Empirical testing against simpler problems or those with smaller gaps, as cited in Ref [57], demonstrated significant improvement over other dual and primal methods, and is thus a promising approach for the problem formulated in Section 2.

In our approach, we relax the coupling constraints that link different trucks, including demand [24], [25] and charging station capacity constraints (26), and penalize their violations. By doing so, we will be able to formulate truck subproblems, thereby improving computational efficiency. This type of relaxation is coupled with penalties for constraint violations that promote the accelerated reduction of constraint violations.

The optimal coordination of trucks is tantamount to the optimization of the nonconvex dual function, which, in a general form can be written as:

\[
\max_{\Lambda} \{q_{\rho}(\Lambda) : \Lambda \in \mathbb{R}^{\lvert \mathcal{N} \rvert \cdot \lvert \mathcal{P} \rvert} \times \mathbb{R}^{\lvert \mathcal{N} \rvert \cdot \lvert \mathcal{P} \rvert} \times \mathbb{R}^{\lvert \mathcal{N} \rvert \cdot \lvert \mathcal{P} \rvert} \},
\]

with

\[
q_{\rho}(\Lambda) = \min_{x, a, d, u, s} \left\{ L_{\rho}(x, a, d, u, s, \Lambda) \right\},
\]

where \( L_{\rho}(x, a, d, u, s, \Lambda) \equiv \sum_{v \in \mathcal{V}} O_v(d, k, x) + O_{pr}(tard) + \Lambda \cdot H(x) + \rho \cdot \| H(x) \|_1 \) is the “absolute-value” Lagrangian function [60]. Here \( \Lambda \) is the vector of Lagrange multipliers, variables \( x \) above collectively define all the binary decision variables; variables \( a, d, u, s \) define the arrival time, departure time, unloading time, and the state of charge decision variables, respectively.

Minimization within (32) is referred to as the “relaxed problem.” \( H(x) \) is a vector of constraint violations defined as:

\[
H(x) = \left( \begin{array}{c}
\sum_{t \in \mathcal{T}_{\text{inbound}}, v \in \mathcal{V}} x_{v,t} \cdot t_{\text{ld}} - \sum_{\mathcal{P}_{\text{nport}}, \mathcal{P}_{\text{nport}}} D_{\text{nport}, \text{nport}} \\
\sum_{t \in \mathcal{T}_{\text{outbound}}, v \in \mathcal{V}} x_{v,t} \cdot t_{\text{ld}} - \sum_{\mathcal{P}_{\text{ndepot}}, \mathcal{P}_{\text{ndepot}}} D_{\text{ndepot}, \text{ndepot}} \\
\sum_{v \in \mathcal{V}, t \in \mathcal{T}} x_{v,t} \cdot \text{chrg}_{v,n,\text{chrg},p} + s_{\text{d} \cdot \text{chrg},p} - C_{\text{n} \cdot \text{chrg}}
\end{array} \right),
\]

where...
where \( sl \) are slack variables used to convert charging station capacity constraints into equality constraints.

The success of Lagrangian Relaxation methods generally depends on the efficiency of the optimization of the non-smooth dual function [31].

To maximize the dual function (refer to Equation (31)), the SLBLR approach updates the Lagrangian multipliers \( \Lambda \) by taking steps \( \alpha^k \) along “surrogate” subgradient directions \( H(\tilde{x}^k) \) (rather than conventional subgradients):

\[
\Lambda^{k+1} = \Lambda^k + \alpha^k \cdot H(\tilde{x}^k).
\]

The components of \( \Lambda \) associated with “charging capacity” constraints are projected onto a positive orthant \( \{ \lambda : \lambda \geq 0 \} \).

To set stepsizes, SLBLR utilizes a decision-driven scheme (see (36) ahead) to determine a series of “level values” that are overestimates of \( q(\Lambda^*) \). Following [37], the step-sizes are set as by using “level values” \( \bar{\nu}_j \) as:

\[
\alpha^k = \zeta \cdot \frac{1}{V} \cdot \frac{\bar{\nu}_j - \rho L(\tilde{x}^k, \tilde{d}^k, \tilde{u}^k, \tilde{s}^k, \Lambda^k)}{\|H(\tilde{x}^k)\|^2}, \quad \zeta < 1.
\]

(35)

Here “tilde” is used to distinguish optimal solutions to the relaxed problem, from the solutions to the relaxed problem obtained subject to the “surrogate optimality condition” (see (43) ahead). Accordingly, \( L_p(\tilde{x}^k, \tilde{d}^k, \tilde{u}^k, \tilde{s}^k, \Lambda^k) \) is referred to as a “surrogate dual value” which is an optimized (not necessarily optimal) value of the “absolute-value” Lagrangian function.

To operationalize the above scheme, the following is required:

1. The values of \( \bar{\nu}_j \) are set after detecting divergence of multipliers through the following “multiplier-divergence-detection” feasibility problem:

\[
\begin{align*}
\|\Lambda - \Lambda^{k+1}\| & \leq \|\Lambda - \Lambda^k\|, \\
\|\Lambda - \Lambda^{k+2}\| & \leq \|\Lambda - \Lambda^{k+1}\|, \\
\vdots \\
\|\Lambda - \Lambda^{k+n_j}\| & \leq \|\Lambda - \Lambda^{k+n_j-1}\|.
\end{align*}
\]

(36)

If the above system of equations admits no feasible solution with respect to \( \Lambda \) for some \( k_j \) and \( n_j \), then \( \exists \kappa \in [k_j, k_j + n_j] \) such that

\[
\bar{\nu}_{\kappa,j} = \max_{\kappa \in [k_j, k_j + n_j]} \bar{\nu}_{\kappa,j} > q_\rho(\Lambda^*).
\]

(37)

where

\[
\bar{\nu}_{\kappa,j} = \alpha^\kappa \cdot V \cdot \|H(\tilde{x}^\kappa)\|^2 + L_p(\tilde{x}^\kappa, \tilde{d}^\kappa, \tilde{u}^\kappa, \tilde{s}^\kappa, \Lambda^\kappa).
\]

(38)

While \( k \) is the iteration number, \( j \) is the “level-value” update number; the same level value \( \bar{\nu}_j \) is used for multiplier iterations \( \kappa \in [k_j, k_j + n_j] \). Iterations \( k_j \) and \( n_j \) are determined post factum based on how long it takes to detect multiplier divergence and how often.

2. The “tilde” is used to denote solutions to the relaxed problem obtained by solving one subproblem at a time rather than carrying out exact minimization within [32]. Subproblems are formulated as follows:

\[
\min \{ O_s(d, \bar{n}, x) + O^{k-1}_p (tard) + \Lambda \cdot H^{k-1}(x) + \rho \cdot \|H^{k-1}(x)\|_1 \},
\]

(39)

where

\[
H^{k-1}(x) = \left( \begin{array}{c}
\sum_{t \in T^{inbnd} \cap T, \nu' \in \Omega_{(pr_{nport}, ndepot)}^{(pr_{nport}, ndepot)} \setminus \{v\} \} x^{ld, k-1}_{t, t'} + x^{ld}_{t, t'} - \sum_{pr_{nport} \in PR} D_{nport, pr_{nport}} \\
\sum_{t \in T^{inbnd} \cap T, \nu' \in \Omega_{(pr_{ndepot}, nport)} \setminus \{v\} \} x^{ld, k-1}_{t, t'} + x^{ld}_{t, t'} - \sum_{pr_{ndepot} \in PR} D_{ndepot, pr_{ndepot}} \\
\sum_{t \in T, \nu' \in \{v\} \} x_{t, t'}^{chrg, k-1} + x_{t, t'}^{chrg} + \sum_{t \in T, \nu' \in \{v\} \} x_{t, t'}^{chrg, k-1} + x_{t, t'}^{chrg} + \sum_{t \in T, \nu' \in \{v\} \} x_{t, t'}^{chrg, k-1} + x_{t, t'}^{chrg} + \sum_{t \in T, \nu' \in \{v\} \} x_{t, t'}^{chrg, k-1} + x_{t, t'}^{chrg}
\end{array} \right)
\]

is a vector of constraint violations and

\[
H(x) = \left( \begin{array}{c}
\sum_{t \in T^{inbnd} \cap T, \nu' \in \Omega_{(pr_{nport}, ndepot)}^{(pr_{nport}, ndepot)} \setminus \{v\} \} x^{ld, k-1}_{t, t'} + x^{ld}_{t, t'} - \sum_{pr_{nport} \in PR} D_{nport, pr_{nport}} \\
\sum_{t \in T^{inbnd} \cap T, \nu' \in \Omega_{(pr_{ndepot}, nport)} \setminus \{v\} \} x^{ld, k-1}_{t, t'} + x^{ld}_{t, t'} - \sum_{pr_{ndepot} \in PR} D_{ndepot, pr_{ndepot}} \\
\sum_{t \in T, \nu' \in \{v\} \} x_{t, t'}^{chrg, k-1} + x_{t, t'}^{chrg} + \sum_{t \in T, \nu' \in \{v\} \} x_{t, t'}^{chrg, k-1} + x_{t, t'}^{chrg} + \sum_{t \in T, \nu' \in \{v\} \} x_{t, t'}^{chrg, k-1} + x_{t, t'}^{chrg}
\end{array} \right)
\]

(40)

(41)
is a vector of surrogate subgradient multiplier-updating directions computed after solving subproblem \textup{(39)}. Moreover, \( O_{pr}^{k-1}(\text{tard}) \), which is defined through tardiness \( tard_{pr} \), is affected by the unloading constraints as follows:

\[
\begin{align*}
\pi_{pr} & \geq u_{v,t,pr}, \forall (t \in T, pr \in \mathcal{P}R), \\
\pi_{pr} & \geq u_{r,t,pr}^{k-1}, \forall (v' \in \mathcal{V} \setminus \{v\}, t \in T, pr \in \mathcal{P}R).
\end{align*}
\]

The difference between \textup{(27)} and \textup{(42)} is that within \textup{(42)}, decision variables associated with electric trucks other than \( v \) are fixed at the latest values available up to the previous iteration \( k - 1 \).

The minimization within \textup{(39)} involves piecewise linear penalties (\( l_1 \) norms), that efficiently penalize constraint violations and are exactly linearizable through the standard use of \textit{special ordered sets}, thereby enabling the use of MILP solvers.

\textbf{Surrogate Optimization}. Even though the decomposition into truck subproblems exponentially reduces computational effort, the effort required to obtain the exact minima of the subproblems \textup{(39)} can be significant. To further reduce the effort, optimization of \textup{(39)} is subject to the following “surrogate optimality condition,” which is an extension of that within \textup{(57)} eq. (11):

\[
\begin{align*}
O_v(\tilde{d}^k, H^k, \tilde{x}^k) + O_{pr}^{k-1}(\text{tard}^k) + \Lambda^k \cdot H^{k-1}(\tilde{x}^k) + \rho^{k} \cdot \|H^{k-1}(\tilde{x}^k)\|_1 < \\
O_v(\tilde{d}^{k-1}, \tilde{d}^{k-1}, \tilde{a}^{k-1}, \tilde{a}^{k-1}) + O_{pr}^{k-1}(\text{tard}^{k-1}) + \Lambda^k \cdot H^{k-1}(\tilde{x}^{k-1}) + \rho^{k} \cdot \|H^{k-1}(\tilde{x}^{k-1})\|_1.
\end{align*}
\]

Essentially, we only seek a solution that improves upon the subproblem incumbent solution, as found up until the previous interaction \( k - 1 \), but taking into account the updated values of multipliers \( \Lambda^k \) and penalty coefficients \( \rho^k \). This approach significantly further reduces the associated computational effort.

\textbf{Algorithm}. The entire algorithm is summarized as follows:

\begin{enumerate}
\item \textbf{Input} \( \Lambda^0, \rho^0, \beta > 1, \zeta < 1, v = 1, k = 1, q_0 = +\infty, q^{max} = -\infty \)
\item \textbf{While} stopping criteria are not satisfied \textbf{do}
\item \textbf{Solve} subproblem \textup{(39)},
\item \textbf{Calculate} \( H(\tilde{x}^k) \),
\item \textbf{Calculate} \( L_\rho(\tilde{x}^k, \tilde{a}^k, \tilde{d}^k, \tilde{u}^k, \tilde{s}^k, \Lambda^k) \),
\item \textbf{Calculate} stepsizes per \textup{(35)},
\item \textbf{Update} multipliers per \textup{(34)},
\item \textbf{If} \( q^{max} < q^{max} \left( \alpha^k \cdot V \cdot \|H(\tilde{x}^k)\|^2 + L_\rho(\tilde{x}^k, \tilde{a}^k, \tilde{d}^k, \tilde{u}^k, \tilde{s}^k, \Lambda^k) \right) \) \textbf{then} \( q^{max} = \alpha^k \cdot V \cdot \|H(\tilde{x}^k)\|^2 + L_\rho(\tilde{x}^k, \tilde{a}^k, \tilde{d}^k, \tilde{u}^k, \tilde{s}^k, \Lambda^k) \)
\item \textbf{EndIf}
\item \( v \leftarrow v + 1 \)
\item \( k \leftarrow k + 1 \)
\item \textbf{If} \( v = V \) \textbf{then} \( v = 1 \)
\item \textbf{EndIf}
\item \textbf{If} \textup{(36)} \textbf{is infeasible} \textbf{then} \( q_j = q^{max}, q^{max} = -\infty, j \leftarrow j + 1 \)
\item \textbf{EndIf}
\item \textbf{If} \( H(\tilde{x}^k) = 0 \) \textbf{then} record feasible cost and \( \rho^k \leftarrow \rho^{k-1} / \beta, \beta > 1. \)
\item \textbf{EndWhile}
\end{enumerate}

\textbf{Brief Discussion of the Algorithm}. In 1, the following criteria can be used: number of iterations, CPU time, or a duality gap. To obtain a duality gap, a feasible cost and a dual value are needed. The feasible cost is obtained per 12, whereas to obtain a dual value, additional effort is needed: the penalty coefficient \( \rho \) needs to be set to 0, and all electric-truck subproblems need to be solved to optimality.
4 Numerical Testing

The solution methodology is implemented in CPLEX 22.1.0.0 with default settings on an Intel(R) Core(TM) i9-9900X 3.50 GHz server with 64 GB RAM. In Example 1, a small instance with 5 trucks, 1 port, and 1 warehouse is considered. The purpose is to compare the performance of the new method with the performance of CPLEX as well as to test the stability of the method. In Example 2, a medium-sized instance with 15 trucks, 1 port, and 1 warehouse is considered. The purpose is to demonstrate that JRC leads to much reduced operational costs as compared to scheduling that considers shortest paths. In Example 3, a large-scale real-world problem in the Greater Los Angeles area with 50 trucks, 1 port, and 3 depots is considered to demonstrate the scalability of the method. Furthermore, Example 3 examines various case studies with different parameters including the number of charging stations, charging speed, and battery capacity size for electric trucks to demonstrate the economic impacts of these variables.

4.1 Example 1: Small Cases with up to 5 Trucks.

This example aims to: 1. Show how combinatorial complexity increases with more trucks, trips, and the units of demand; 2. Reveal that an increase in trucks to five renders conventional solvers unable to find feasible solutions; and 3. Confirm the SLBLR method’s stability. We compare SLBLR to standard Lagrangian Relaxation, which updates multipliers via subgradient directions and adjusts step sizes based on levels \[58\], to ensure a fair assessment.

In this example, a transportation network with 5 nodes (shown in Figure 2) is considered. Road segments are labeled with travel times, assumed equal in both directions. The charging rate is 17% per period, with discharge rates at 5% when loaded and 2.5% when not. Charging station capacities at each node are shown next to the respective road segments.

The first case contains one truck having to transport one unit of cargo to the port and one unit of cargo to the depot on the way back. As demonstrated in Table 2 (first row), this case is solved within a matter of a couple of seconds. As the demand and the number of trips double (second row), the CPU time is 130.38 seconds - a 58-time increase. As the demand and the number of trucks double (third row), the CPU time increases to 70,134.89 seconds, thereby constituting another 537-fold increase over the results presented in the second row. This is a signature behavior of combinatorial NP-hard discrete optimization problems - as the size of the problem increases linearly, the resulting complexity increases superlinearly (e.g., exponentially).

Table 2: Illustration of the complexity increase with the increase of the demand, the number of trips, and the number of trucks.

<table>
<thead>
<tr>
<th>$v$</th>
<th>$t$</th>
<th>Demand at $n_{port}$</th>
<th>Demand at $n_{depot}$</th>
<th>Constraints</th>
<th>Binaries</th>
<th>Nonzeros</th>
<th>Time periods</th>
<th>CPU Time (Sec)</th>
<th>Gap (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>10,485</td>
<td>2,958</td>
<td>289,418</td>
<td>65</td>
<td>2.23</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>18,474</td>
<td>5,447</td>
<td>478,529</td>
<td>65</td>
<td>130.38</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>36,060</td>
<td>10,680</td>
<td>943,043</td>
<td>65</td>
<td>70,134.89</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>3</td>
<td>8</td>
<td>151,791</td>
<td>28,312</td>
<td>4,782,805</td>
<td>65</td>
<td>1,000,000</td>
<td>No sln. found</td>
</tr>
</tbody>
</table>

Ultimately, consider the case where five trucks are assigned to manage both 3 and 8 units of demand at the port and depot respectively. This logistics scenario requires a two-step transportation process. Initially, three units of cargo need to be transported from Node 1 (the depot) to Node 5 (the port) for the purpose of export. Following the export process, there is a requirement to pick up eight units of cargo from Node 5 and transport them back to Node 1 for import. Despite the significant computational effort, the problem remains unsolved after 1,000,000 seconds of CPU time.

Figure 2: Network topology for Examples 1 and 2. The number of time periods required to travel (in both directions) is shown next to the respective road segments.
As described in subsection 2.5, the number of trips for the original Example 1 (which is also reported in the last row of the table above) is estimated to be 4 and the longest path from a depot to a port is estimated to be 15 \((1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5)\). Therefore, the number of periods is 60. An additional 5 units of time is heuristically added as a cushion. For other newly added scenarios, the number of time periods is kept the same while varying other parameters like the number of trips, vehicles, and demand/supply in order to see how these parameters affect the complexity.

Figure 3 demonstrates the performance of the SLBLR method. The first feasible solution found by the SLBLR method is obtained within less than 100 seconds. The method intrinsically possesses the ability to systematically improve the solution, since the “price-based” and systematical update of the Lagrangian multipliers leads to “more economical” solutions. Moreover, subproblem solutions, when close to feasibility, are easily repaired to obtain feasible solutions. As a result, the method is capable of improving upon previously obtained solutions with ease. Ultimately, a high-quality solution with a gap of less than 1% is obtained within 5 minutes.

Comparison against other methods. We compare our method against standard package CPLEX and standard Lagrangian Relaxation (LR) with subgradient directions used to update multipliers.

1. **Comparison against CPLEX**: In attempting to solve the same problem instance, CPLEX is unsuccessful in finding a feasible solution even after 1,000,000 seconds. This puts SLBLR ahead by 4 orders of magnitude in terms of speed. When analyzing lower bounds and solution quality, CPLEX could only compute a lower bound of 3.1452 within the time limit, failing to find a feasible solution. This implies that even with the help of cutting and branching, CPLEX could not make a significant impact on improving the lower bound. It is essential to note that while there is a significant difference between the CPLEX lower bound and ours, we attribute this to the distinct methodologies and heuristics each system employs. As a result, these varying approaches can lead to different lower bounds and overall performances.

![Figure 3: Solutions obtained by the SLBLR method for Example 1](image)

2. **Comparison against LR with Subgradient-Level approach**\(^7\) to update multipliers: While LR’s strength is the exploitation of the drastic reduction of complexity upon decomposition, standard LR requires that 1. truck subproblems are solved to optimality and 2. all truck subproblems are solved before updating multipliers. In contrast, our method requires subproblem-feasible solutions that satisfy the “surrogate optimality condition” which only requires that only one solution that can improve the subproblem incumbent is obtained. Our method requires only 8 seconds to solve a truck subproblem, whereas standard LR requires 566 seconds. After 22,656.31 seconds, LR obtained a feasible solution with a cost of 24 and a lower bound of 18.6 (which is a gap of 22.5%), whereas our method found a solution with a cost of 20.7 and a lower bound of 20.6 (which is a gap of 0.5%) after 287 seconds. The results are shown in Figure 4. While the standard Lagrangian Relaxation is more successful than CPLEX, which does not exploit the reduction of complexity upon decomposition, our method is faster than standard LR by roughly 2 orders of magnitude.

\(^7\)Subgradient-Level method\(^5\) is chosen because, like our method, it also uses the Polyak formula\(^6\), albeit, with a different technique to obtain level values.
Figure 4: Trajectories of multipliers within Surrogate “Level-Based” Lagrangian and standard Lagrangian Relaxation with Subgradient-Level method for updating multipliers.

Figure 4 demonstrates trajectories of multipliers generated by standard LR and our method. Within standard LR, multipliers zigzag whereas directions obtained by our method are much smoother. More importantly, our method requires much less time per multipliers update as discussed above.

**Stability of SLBLR.** The stability with respect the coordination-related hyperparameter $\zeta$ was previously established in the foundational study detailed in [57]. Consistent with this, we have maintained $\zeta$ at a constant value of $1/2$ throughout all Examples. However, the original investigation [57] did not explore the stability of the primal-solution-recovery parameters, namely $\beta$ and $\rho_0$. In this context, we find it imperative to extend the stability analysis to these parameters and, hence, we present our findings herein.

For $\rho_0$ and $\beta$, the baseline values used to obtain results above are 5 and 1.05 respectively. The results of the experiments when varying these parameters are provided in the subsequent tables:

**Table 3: CPU times for varying values of $\beta$.**

<table>
<thead>
<tr>
<th>$\rho_0$</th>
<th>$\beta$</th>
<th>CPU Time (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>1.1</td>
<td>269</td>
</tr>
<tr>
<td>5</td>
<td>1.05</td>
<td>287</td>
</tr>
<tr>
<td>5</td>
<td>1.025</td>
<td>420</td>
</tr>
</tbody>
</table>

**Table 4: CPU times for varying values of $\rho_0$.**

<table>
<thead>
<tr>
<th>$\rho_0$</th>
<th>$\beta$</th>
<th>CPU Time (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>1.05</td>
<td>313</td>
</tr>
<tr>
<td>5</td>
<td>1.05</td>
<td>287</td>
</tr>
<tr>
<td>2.5</td>
<td>1.05</td>
<td>314</td>
</tr>
</tbody>
</table>

According to Tables 3 and 4 results are fairly stable. The computational performance was evaluated based on the time required to achieve a solution with a cost of 20.7—the near-optimal result achieved in our experiments as explained above.

Stability concerning initialization procedures was addressed in the seminal SLBLR paper [57]. Recognizing that initialization stability may vary with the nature of the problem, our approach included a stability test via randomized multiplier initialization. Multipliers were uniformly distributed within the interval $[-50, 50]$. The CPU times for this process are summarized in Table 5.
Table 5: CPU times for randomized initialization of multipliers.

<table>
<thead>
<tr>
<th>Test Instance</th>
<th>CPU Time (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>446</td>
</tr>
<tr>
<td>2</td>
<td>472</td>
</tr>
<tr>
<td>3</td>
<td>399</td>
</tr>
<tr>
<td>4</td>
<td>324</td>
</tr>
<tr>
<td>5</td>
<td>308</td>
</tr>
<tr>
<td>6</td>
<td>362</td>
</tr>
<tr>
<td>7</td>
<td>524</td>
</tr>
<tr>
<td>8</td>
<td>302</td>
</tr>
<tr>
<td>9</td>
<td>331</td>
</tr>
<tr>
<td>10</td>
<td>251</td>
</tr>
</tbody>
</table>

| Average       | 372            |
| Standard Deviation | 86             |

As compared to a CPU time obtained to solve the instance of Example 1 whereby the multipliers are all initialized at 0, which is 287 seconds, the above average time is 29.72% since some of the random multipliers may fall far from the actual optimum. Overall, given the random nature of this testing, the increase in the CPU time is rather modest.

Figure 5: A side-by-side comparison of schedules obtained by using the novel modeling versus the “shortest path” approach for Example 2.

4.2 Example 2: Medium-Scale Case with 15 Trucks.

In this example, the same road network is shown in Figure 2 as well as charging station capacities and due times. The goal of a fleet of 15 trucks is to deliver (to export) 8 units of cargo from node 1 (Depot) to node 5 (Port) and to pick up (to import) 24 units of cargo from node 5 and deliver them to node 1. In order to provide insights into the advantages of joint electric truck routing and charging, results obtained by using the novel formulation are compared with the results obtained while considering shortest paths by disregarding nodes 2 and 4, since visiting these nodes increases the distance and the time traveled. For both cases, the CPU time limit was set to 10 minutes; for the joint scheduling the cost is 68.2 and for the “shortest path” scheduling the cost is 73.4. The scheduling results are demonstrated in
Figure 5. Within JRC, since the trucks can take detours through nodes 2 and 4, they save time needed to wait for the availability of the charging station. Ultimately, the overall time (and tardiness) decreases.

For example, trucks 1 and 3 (Figure 5 (top)) take a detour through node 4 to charge, and trucks 4 and 5 take a detour through node 2 since chargers at node 3 are occupied by trucks 8 and 11. In contrast, truck 1 (Figure 5 (bottom)) needs to wait 6 time periods during the first round trip to charge at node 3 because the charger is successively occupied by trucks 4, 7, and 5. Likewise, truck 2 needs to wait three time periods during both the first and the second round trips because the charger at node 3 is occupied by trucks 6 and 8 (during the first trip) and by trucks 3 and 4 (during the second trip). Overall, not only does joint routing and charging lead to a decrease in the cost but also one truck is spared (truck 15). We also discovered that the total time trucks wait for the available charger is reduced by 69%, and the overall operational cost is reduced by 6.8%.

4.3 Example 3: Large-Scale Cases with 50 Trucks in the Greater Los Angeles Area.

In this example, a realistic road network topology within the Greater Los Angeles area with one port located at Long Beach and three warehouses located at node 4 (intersection of I-405 and SR 60), node 6 (intersection of SRs 57 and 60), and node 7 (intersections of SRs 60 and 91 with I-215) is adopted as shown in Figure 6.

Figure 6: Road network, port, warehouses, and charging stations for Example 3.

A fleet of 50 trucks operates from warehouses at nodes 4, 6, and 7, containing 17, 13, and 20 trucks, respectively. The trucks are equipped with battery sizes of either 250 kWh [12,61] or 600 kWh (estimated from [62]), featuring standard and long-range versions. The charging power of the base case is assumed to be 350 kW [63]. We follow the regression model in [12] to determine energy consumption rates of fully-loaded and empty trucks. Assuming that the weight of a fully loaded and an empty truck is 80,000 lb and 22,000 lb [61], respectively, we can obtain energy consumption rates of 2.267 kWh/mile for a fully loaded truck and 1.617 kWh/mile for an empty truck. We will study the impact of higher charging powers considering the industry trends [64]. The import requirements are 39, 35, and 33, and the export requirements are 32, 20, and 25, respectively. A number of testing cases are considered in this example.

**Base Case:** In this case, 30 trucks with 600 kWh batteries are located at the City of Industry and Riverside, and 20 trucks with 250 kWh batteries are located at El Monte. There are 3 chargers available at each node with a maximum charging power of 350 kW. The electric trucks associated with depots 6 and 7 have a longer range with a battery capacity of 600 kWh, while the electric trucks associated with node 4 have a smaller battery capacity of 250 kWh.

**Cases 1.1-1.7:** In this series of cases, we test the impacts of the number of chargers on the total operation cost. Accordingly, the number of chargers at nodes 1-7 is reduced from 3 to 1.

**Case 2:** In this case study, we evaluate the impacts of the size of the electric truck battery on the total operation cost. It is assumed that all trucks are equipped with a battery capacity of 600 kWh.

**Case 3:** In this case study, we quantify the impacts of maximum charging power on the total operation cost. The maximum charging power of all chargers is increased from 350 kW to 700 kW.

The optimization stopping criterion is 1800 seconds. The results for all the case studies are reported in Table 6.
When the number of chargers is reduced from 3 to 1, the overall operation cost increases the most in Cases 1.2 (8.6%), 1.3 (7.8%), 1.1 (7.3%), and 1.4 (6.7%) compared to the Base case. This means that the ordering of the best locations for electric truck charging stations is node 2 (intersection of I-710 and SR 60), node 3 (intersection of I-605 and SR 91), node 1 (Port of Long Beach), and node 4 (intersection of I-405 and SR 60). When we compare Case 2 with the Base Case, the increase in battery capacity for the 20 trucks at El Monte from 250 kWh to 600 kWh leads to a reduction in the total operation cost by 27.9%. Additionally, this enhancement reduces the requirement for the number of electric trucks from 49 to 43. Finally, comparing Case 3 and the Base case, we discover that increasing the maximum charging power of all chargers from 350 kW to 700 kW decreases the operation cost by 9.2%.

Table 6: The Feasible Operation Cost and CPU Time of Test Cases 1-3.

<table>
<thead>
<tr>
<th>Case</th>
<th>Feasible Cost</th>
<th>CPU Time (Sec)</th>
<th>Number of Trucks Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>5527.39</td>
<td>894.53</td>
<td>49</td>
</tr>
<tr>
<td>Case 1.1</td>
<td>5934.63</td>
<td>1151.99</td>
<td>49</td>
</tr>
<tr>
<td>Case 1.2</td>
<td>6005.08</td>
<td>1149.49</td>
<td>49</td>
</tr>
<tr>
<td>Case 1.3</td>
<td>5959.69</td>
<td>887.32</td>
<td>49</td>
</tr>
<tr>
<td>Case 1.4</td>
<td>5900.92</td>
<td>929.83</td>
<td>49</td>
</tr>
<tr>
<td>Case 1.5</td>
<td>5578.95</td>
<td>1150.93</td>
<td>49</td>
</tr>
<tr>
<td>Case 1.6</td>
<td>5683.47</td>
<td>1662.93</td>
<td>49</td>
</tr>
<tr>
<td>Case 1.7</td>
<td>5574.17</td>
<td>1569.50</td>
<td>49</td>
</tr>
<tr>
<td>Case 2</td>
<td>4020.51</td>
<td>1153.15</td>
<td>43</td>
</tr>
<tr>
<td>Case 3</td>
<td>5017.30</td>
<td>1327.38</td>
<td>48</td>
</tr>
</tbody>
</table>

Figure 7: Total cost (y-axis) dependence on truck battery capacity. CPU times (in seconds) are also shown next to data points.

Analysis of the Impacts of Battery Capacity and Maximum Charging Power on the Cost and CPU Time.

In this part of the Example, a network illustrated in Figure 6 is considered, and a fleet of 50 trucks with 250 kWh batteries is assumed. The fleet operates across nodes 4, 6, and 7, housing 10, 20, and 20 trucks respectively. The base case assumes a charging capacity of 350 kW with six chargers per node, import demands of 29, 21, and 15, as well as export demands of 25, 23, and 14, with due times set at 22 and 9 time periods respectively.

The analysis explores the influence of varying battery capacities and charging powers on operational costs and solving times. Fixing the charging power at 110% (385 kW) of its nominal value, battery capacities of 110%, 130%, 150%,
and 170% (275 kWh, 325 kWh, 375 kWh, and 425 kWh) are considered. Findings, depicted in Figure 7, indicate negligible cost benefits beyond a 170% battery capacity increase.

Similarly, with batteries set at 150% (375 kWh) of nominal capacity, charging powers of 110%, 130%, 150%, and 170% (385 kW, 455 kW, 525 kW, and 595 kW) are considered. Results (Figure 8) suggest that increasing charging power above 170% yields no substantial cost advantage. Notably, enhanced battery capacity and charging power correlate with reduced problem-solving time, as longer-range trucks and higher-powered chargers lead to fewer charging events and thus lower the computational complexity.

Figure 8: Total cost (y-axis) dependence on the maximum charging power of charging stations. CPU times (in seconds) are also shown next to data points.

Discussion on Solution Quality. Table 7 demonstrates the computation time required to obtain lower bounds to reach the gap of at least 1% for the feasible costs for five different cases that are demonstrated in Figure 8.

<table>
<thead>
<tr>
<th>Maximum Charging Power Increase</th>
<th>Time (Sec)</th>
<th>Gap (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>111,140</td>
<td>1</td>
</tr>
<tr>
<td>1.3</td>
<td>104,752</td>
<td>1</td>
</tr>
<tr>
<td>1.5</td>
<td>101,110</td>
<td>1</td>
</tr>
<tr>
<td>1.7</td>
<td>33,774</td>
<td>1</td>
</tr>
<tr>
<td>1.9</td>
<td>33,027</td>
<td>0.13</td>
</tr>
</tbody>
</table>

While feasible costs are quickly obtained due to the four advantages highlighted in the Introduction (decomposition, surrogate optimization, rapid coordination, and penalization), computing lower bounds for dual values, which provide lower bounds, is more complex. Exact optimality in truck subproblems is necessary for these calculations. The benefits of surrogate optimization for obtaining primal solutions do not extend to lower bound calculations, leading to longer CPU times for solution quality verification. Generally, achieving a satisfactory lower bound quickly is challenging across various problems, with few exceptions. The reduction of CPU time for robust lower bound computation remains a significant challenge and is beyond the scope of this paper.

Scalability with Respect to the Number of Nodes. The above results demonstrate scalability with respect to the number of trucks. Because of the NP-hardness, upon the decomposition into truck subproblems, the complexity decreases superlinearly. Each individual truck subproblem, however, is an NP-hard problem prone to a combinatorial increase in complexity unless the problem is further decomposed into nodal subproblems. A more granular decomposition is outside the scope of the paper. Without the abovementioned decomposition, we test the limits of the subproblem-solving procedure by using off-the-shelf software. We consider the same network as before but with the addition of two nodes
- 8 and 9 (Figure 9). Our findings are that while the average running time for the 5 cases in Figure 8 stood at 954 seconds, with the addition of these two nodes, the running time increased significantly to 4938 seconds.

![Figure 9: An augmented road network, port, warehouses, and charging stations for Example 3.](image)

### 5 Conclusion

This paper develops a novel formulation of the joint routing and charging (JRC) problem for heavy-duty electric trucks. The optimization problem is to minimize total tardiness as well as transportation and charging costs. Through the application of decomposition and coordination principles, the proposed solution methodology is computationally efficient, allowing solving of the associated discrete optimization problem quickly. The newly developed method can obtain near-optimal solutions within a few minutes for small cases and within 30 minutes for large ones. Furthermore, it has been demonstrated that as truck battery size increases, the total cost decreases significantly; moreover, as maximum charging power increases, the number of electric trucks required decreases as well.

**Broader Impact.** In addition to its ability to drastically reduce the complexity of the JRC problem, the proposed method also provides Lagrangian multipliers, which can be treated as “shadow prices” and used to provide intuitive explanations of the underlying decision-making process. The solution methodology can also run “reactively” and online to adapt to unexpected events such as truck/charging station breakdowns, blockage of road segments due to traffic accidents, etc. In the presence of such events, Lagrangian multipliers are readjusted “on-the-fly” thereby providing re-routing guidance for the trucks. Since reducing wait times upstream would reduce propagation delays downstream, a lower tardiness rate would be beneficial for supply chain and downstream operations management.

**Future Directions.** As highlighted in [65], drivers may adjust speed to reduce energy consumption. Therefore, additional functionality including a choice of speed will be added to the novel formulation. Furthermore, since there may be uncertainties regarding demand [66, 67], availability of charging stations [4], and arrival times for electric trucks [68], the stochastic joint routing-charging problem (JRC) will also be considered. Although the above considerations add additional complexity to the solving process, preliminary results of a study in [69] indicate that well-trained deep neural networks can generate subproblem solutions within a matter of milliseconds.

### Nomenclature

**Superscripts**

- **arrvl** Indicates “arrival” for decision variables capturing truck arrival
- **bgn** Indicates “begin” for decisions when to begin charging
- **chrg** Indicates “charge” for nodes containing chargers (if used as a superscript of a set), for charging costs and charge rates (if used as a superscript of input data) as well as for a superscript of decision variables related to charging
- **cmplt** Indicates “complete” for decisions when to complete charging
- **depot** Indicates “depot” for depot nodes
Indicates “departure” for decision variables capturing truck departure
Indicates “empty” for input data to capture the discharge rate of a truck
Indicates “inbound” for the set of trips from port to depot
Indicates “load” for binary variables that capture the “load” status of a truck
Indicates “loaded” for input data to capture the discharge rate of a truck
Indicates “outbound” for the set of trips from depot to port
Indicates “port” for port nodes
Indicates “trip” for decision variables capturing trips
Indicates “travel” for travel time

**Parameters**

- \(N\) Number of nodes in the transportation network
- \(P\) Number of time periods
- \(R\) Number of road segments
- \(T\) Number of trips
- \(\beta\) Parameter for updating penalty coefficient to penalize constraint violations
- \(\zeta\) Parameter for updating the Polyak stepsize within SLBLR to ensure appropriate reduction of stepsizes
- \(q^{max}\) Candidate for the level value
- \(V\) Number of vehicles - trucks

**Sets**

- \(\mathcal{N}\) Set of nodes
- \(\mathcal{N}^{\text{depot}}\) Set of depot nodes
- \(\mathcal{N}^{\text{port}}\) Set of port nodes: \(\mathcal{N}^{\text{depot}} \cap \mathcal{N}^{\text{port}} = \emptyset\)
- \(\mathcal{PR}\) Set of products
- \(\mathcal{P}\) Set of time periods
- \(\mathcal{R}\) Set of road segments
- \(\mathcal{T}\) Set of trips
- \(\mathcal{V}\) Set of vehicles - trucks

**Indices**

- \(n\) Node number \(n \in \mathcal{N} = \{1, \ldots, N\}\)
- \(n^{\text{depot}}\) Depot node number \(n^{\text{depot}} \in \mathcal{N}^{\text{depot}} (\neq \emptyset) \subset \mathcal{N}\)
- \(n^{\text{port}}\) Port node number \(n^{\text{port}} \in \mathcal{N}^{\text{port}} (\neq \emptyset) \subset \mathcal{N}\)
- \(p\) Time period \(p \in \mathcal{P} = \{1, \ldots, P\}\)
- \(r\) Road segment \(r \in \mathcal{R} = \{1, \ldots, R\}\)
- \(s(r)\) Starting node of road segment \(r\)
- \(e(r)\) Ending node of road segment \(r\)
- \(t\) Trip number \(t \in \mathcal{T} = \{1, \ldots, T\}\)
- \(j\) Level value update number
- \(k\) Iteration number within Lagrangian Relaxation
- \(pr\) Product number
- \(pr_n\) Product that needs to be delivered to node \(n\)
- \(v\) Vehicle number \(v \in \mathcal{V} = \{1, \ldots, V\}\)

**Input Data**

- \(C_{p,n}^{\text{chrg}}\) Charging cost per time period at node \(n^{\text{chrg}}\) is charging starts at time \(p\)
due_{pr}  Due time for product pr
\Delta s^{chrg}  Charging rate (in %)
\Delta s^{dchrg,empty}  Discharge rate of an empty truck (in %)
\Delta s^{dchrg,ldd}  Discharge rate of a loaded truck (in %)
av_{v}  Available time of truck v at the beginning of the scheduling horizon
\text{C}^{lbr}  Labor cost per time period after the first departure and before the last arrival (assumed to be independent of the time of the day as well as the truck)
\text{C}_{n^{chrg}}  Number of chargers at a charging location n^{chrg}
\text{C}^{chrg}_{p,n^{chrg}}  Charging cost per time period if charging starts at time p at node n^{charge}
\text{T}^{trvl}_{r,p}  Travel time through road segment r if a truck departs from s(r) at time p

Decision Variables
\Lambda  Dual continuous variables (Lagrangian multipliers)
\pi_{v,v^{depot}}  Integer variables to capture the latest arrival time of vehicle v at the depot v^{depot}
\pi_{pr}  Integer variable to capture the latest unloading time for product pr
\alpha_{v,n,t}  Integer variable for arrival time of truck v to node n during trip t
\beta_{v,n^{chrg},t}  Integer variable for beginning of the charging time of truck v at a node n^{chrg} equipped with charging stations during trip t
\text{C}^{tard}_{pr}  Tardiness per period penalty for late delivery of product pr
\gamma_{v,n^{chrg},t}  Integer variable for completion of the charging time of truck v at a node n^{chrg} equipped with charging stations during trip t
\eta_{v,n,t}  Integer variable for departure time of truck v from node n during trip t
\text{s}^{chrg}_{v,n,t}  State of charge continuous decision variable (\in [0,1]) of a battery of truck v at node n during trip t
\text{sl}  Vector of slack variables for charging station capacity constraints
\text{tard}_{pr}  Integer variable to capture the tardiness for product pr
\upsilon_{v,n,t,pr}  Integer variable for product pr unloading start time by truck v at node n (which needs to be either a port or a depot) during trip t
\chi^{chrg}_{v,n^{chrg},p,t}  Binary variable to capture whether truck v charges at node n^{chrg} equipped with chargers) at time p during trip t
\chi^{ld}_{v,n,t}  Binary variable to capture whether truck v is loaded during trip t
\chi^{trip}_{v,t}  Binary variable to capture whether trip t is taken by truck v
\chi^{trvl,arrvl}_{v,n,p,t}  Binary variable that captures whether truck v arrived to node n at time p during trip t
\chi^{trvl,dprt}_{v,n,p,t}  Binary variable that captures whether truck v departed from node n at time p during trip t

Functions
\alpha^{k}  Stepizes as a function of iteration number
\tau_{j}  A level value after j\text{th} update
\rho^{k}  Penalty coefficient as a function of iteration number
H(x)  Levels of constraint violations
L_{\rho}(x,a,d,u,s,\Lambda)  Lagrangian function as a function of charging, departure, arrival, and trip decisions collectively denoted by x, arrival time a, departure time d, unloading time u, state of charge s, Lagrangian multipliers \Lambda and penalty coefficient \rho
\text{O}_{v}(d,\pi,x)  A part of the objective function for truck v that includes labor and charging costs
\text{O}_{pr}(\text{tard})  A part of the objective function for truck v that includes tardiness cost
\text{q}_{\rho}(\Lambda)  Dual function as a function of Lagrangian multipliers \Lambda and penalty coefficient \rho
A  Linearization of Logical Constraints

A.1 Linearization of Logical Constraints with Inequalities such as (1).

Logical constraint (1) as well as other logical constraints thereafter are linearized by using big-M constraints as such:

\[
d_{v,n_v} \geq a_v - M \cdot (1 - x_{v,1}^{trip}), \forall \ (v \in \mathcal{V}),
\]

where \( M \) is a large number. Indeed, if \( x_{v,1}^{trip} = 1 \), then \( d_{v,n_v} \geq a_v \) as required by (1); if \( x_{v,1}^{trip} = 0 \), then (1) is redundant.

A.2 Linearization of Logical Constraints with Equalities such as \((5)\) and \((7)\).

Linearization of \((5)\) follows the same principle, with the exception that two sets of constraints are needed. For illustration purposes, linearization of \((5)\) is demonstrated since other constraints are linearized in the same way:

\[
\sum_{r \in \mathcal{R}, s(r) = n} x_{v,e(r),p+T_{r,p}^{trvl},1,t} - 1 \leq M \cdot (1 - x_{v,n,p,t}^{dprtr}), \forall \ (v \in \mathcal{V}, n \in \mathcal{N}, t \in \mathcal{T}, p \in \mathcal{P}),
\]

\[
\sum_{r \in \mathcal{R}, s(r) = n} x_{v,e(r),p+T_{r,p}^{trvl},1,t} - 1 \geq -M \cdot (1 - x_{v,n,p,t}^{dprtr}), \forall \ (v \in \mathcal{V}, n \in \mathcal{N}, t \in \mathcal{T}, p \in \mathcal{P}).
\]

A.3 Linearization of Logical Constraints with Equalities and Several Logical Statements such as \((8)\) and \((10)\).

Linearization of \((8)\) is performed as:

\[
s_{v,e(r),t}^{chrg} = s_{v,n,t}^{chrg} - \frac{\Delta s_{v}^{dchrg,ldd}}{100} \cdot T_{r,p}^{trvl} \cdot x_{v,e(r),p+T_{r,p}^{trvl},1,t} \leq 0,
\]

\[
M \cdot (2 - x_{v,n,p,t}^{dprtr} - x_{v,e(r),p+T_{r,p}^{trvl},1,t}) \leq 0,
\]

\[
\forall \ (v \in \mathcal{V}, n \in \mathcal{N}, r \in \mathcal{R} : s(r) = n, t \in \mathcal{T}, p \in \mathcal{P} : p + T_{r,p}^{trvl} - 1 \leq P).
\]

Accordingly, linearization of \((9)\) is performed as:

\[
s_{v,e(r),t}^{chrg} = s_{v,n,t}^{chrg} - \frac{\Delta s_{v}^{dchrg,ldd}}{100} \cdot T_{r,p}^{trvl} \cdot x_{v,e(r),p+T_{r,p}^{trvl},1,t} \geq 0,
\]

\[
M \cdot (2 - x_{v,n,p,t}^{dprtr} - x_{v,e(r),p+T_{r,p}^{trvl},1,t}) \geq 0,
\]

\[
\forall \ (v \in \mathcal{V}, n \in \mathcal{N}, r \in \mathcal{R} : s(r) = n, t \in \mathcal{T}, p \in \mathcal{P} : p + T_{r,p}^{trvl} - 1 \leq P).
\]

Linearization of other constraints follows the same principles. For example, linearization of \((15)\), \((20)\) and \((21)\) follows subsection \(A.2\) \((16)\), \((17)\), \((18)\) and \((19)\) follows \(A.1\) and \(A.3\).
References


