



**HAL**  
open science

## Smart Charging of Electric Vehicles: an Autonomous Driving Perspective

Fabrizio Sossan, Charitha Buddhika Heendeniya, Biswarup Mukherjee, Vasco Medici

► **To cite this version:**

Fabrizio Sossan, Charitha Buddhika Heendeniya, Biswarup Mukherjee, Vasco Medici. Smart Charging of Electric Vehicles: an Autonomous Driving Perspective. [Research Report] MINES ParisTech - Université PSL; SUPSI. 2022, pp.1-26. hal-03756809

**HAL Id: hal-03756809**

**<https://hal.science/hal-03756809>**

Submitted on 22 Aug 2022

**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.

Copyright



# SMART CHARGING OF ELECTRIC VEHICLES: AN AUTONOMOUS DRIVING PERSPECTIVE

VERSION 1.0

Fabrizio Sossan  
Charitha Buddhika Heendeniya  
Biswarup Mukherjee  
Vasco Medici

**May 2022**

## INTERNAL REFERENCE

<b>Deliverable No.:</b>	D 4.1 (2022)
<b>Deliverable Name:</b>	Advanced charge-scheduling algorithms for EVs and EVAs and forecasting
<b>Lead Participant:</b>	Originally ETHZ, now MINES ParisTech
<b>Work Package No.:</b>	WP4
<b>Task No. &amp; Name:</b>	T4.1 Advanced charge-scheduling algorithms for EVs and EVAs and forecasting
<b>Document (File):</b>	EVA D4.1
<b>Issue (Save) Date:</b>	2022-05-31

## DOCUMENT STATUS

	Date	Person(s)	Organisation
<b>Author(s)</b>	2022-05-30	Charitha Buddhika Heendeniya	SUPSI
	2022-05-30	Vasco Medici	SUPSI
	2022-05-30	Fabrizio Sossan	MINES ParisTech
	2022-05-30	Biswarup Mukherjee	MINES ParisTech
<b>Verification by</b>	2022-07-01	Albedo Bettini	SUPSI
<b>Approval by</b>	2022-07-04	Roman Rudel	SUPSI
<b>Approval by</b>			

## DOCUMENT SENSITIVITY

- Not Sensitive** Contains only factual or background information; contains no new or additional analysis, recommendations or policy-relevant
- Moderately Sensitive** Contains some analysis or interpretation of results; contains no recommendations or policy-relevant statements
- Sensitive** Contains analysis or interpretation of results with policy-relevance and/or recommendations or policy-relevant statements
- Highly Sensitive Confidential** Contains significant analysis or interpretation of results with major policy-relevance or implications, contains extensive recommendations or policy-relevant statements, and/or contain policy-prescriptive statements. This sensitivity requires SB decision.

## Executive Summary

Electrification of the transportation sector has a range of implications related to the environment, climate change, business, and economy. While there is the political will to support the growth of electric vehicles (EV), the charging process of large populations of EVs is challenging because the existing electrical infrastructure, and in particular power distribution grids, might not be capable of meeting increased levels of power demand.

In the existing literature, smart charging algorithms have been widely investigated to resolve the issues related to the simultaneous charge of many EVs. This report takes the perspective of autonomous driving, which is expected to revolutionize road transport. In particular, this report explores how autonomous EVs (AEVs) could facilitate better management of the charging process, thanks to the innovations in data-driven control methods for making smart charging decisions.

We present two architectures for smart charging for AEVs, a model-based controller and a model-free stochastic reinforcement learning (RL) controller, enabling grid-friendly and optimal integration of EVs and AEVs in a smart grid.

The model-based approach leverages an optimal power flow (OPF) to schedule the charging process of AEVs. Compared to OPF-based smart charging algorithms for standard EVs where the charging location is totally determined by the location where the vehicles have been parked by their drivers, AEVs can change location autonomously and choose a more suitable charging spot for the power grid. It is shown how this problem can be formulated as a mixed-integer linear problem (MILP), leveraging linearized load flow models that can be solved with off-the-shelf optimization libraries. It is shown that the proposed scheduler for AEVs can achieve significantly better congestion management than traditional EVs, postponing grid reinforcement.

The second proposed approach uses a stochastic RL controller. The foremost opportunity to use stochastic RL is that a smart-grid environment is typically partially observable. That is to say, due to technical and investment limitations, it is challenging to observe the state of the smart grid in full for control purposes. Moreover, the uncertainties in a smart grid environment affect the controller's optimality.

Designing a stochastic RL controller is a challenging task that requires careful feature engineering, state abstractions, reward engineering, and hyperparameter tuning. However, we can easily roll out the trained stochastic RL agent for optimal charging management in a decentralized control architecture after training. As we show in the case studies, stochastic RL agents effectively find the optimal control policy based on the trade-offs in the reward function under stochastic conditions and imperfect information. Furthermore, the actor-critic architecture presented in the report has the advantage of scalability. For example, the second case study uses an architecture with one critic and many actors. It is also possible to build architectures with many critics and many actors for cases even larger. On the other hand, stochastic RL agents have difficulty handling hard constraints, which are essential for safety-critical applications. However, through proper training methods, it is possible to minimize the chance of constraint violations.

Finally, we see that both model-based and model-free controllers have their advantages and unique place in a smart grid for optimal charging management. From a futuristic viewpoint, it is vital to understand the unique advantages and disadvantages of model-based and model-free controllers, their unique use cases, and even opportunities for hybrid mechanisms that build upon the strengths of both strategies.

## Table of Contents

<b>Executive Summary</b>	<b>2</b>
<b>1 Introduction</b>	<b>2</b>
<b>2 A model-based approach using an augmented optimal power flow</b>	<b>3</b>
2.1 Problem formulation . . . . .	3
2.2 Case study . . . . .	8
2.3 Results . . . . .	10
<b>3 Model-free stochastic control with deep reinforcement learning</b>	<b>10</b>
3.1 A motivating example for partial observability in smart-grid . . . . .	12
3.2 Theoretical background . . . . .	13
3.3 Case study 1 . . . . .	15
3.4 Results - case study 1 . . . . .	16
3.5 Case study 2 . . . . .	18
3.6 Results - case study 2 . . . . .	19
<b>List of publications</b>	<b>22</b>

## 1 Introduction

Electrification of the mobility sector is at the top of the decarbonization agenda for many countries. Several countries have already taken policy steps to either heavily restrict or ban internal combustion vehicles within the next decade [CBER21]. It also enables further innovations in the transportation sector, such as one-way electric car sharing that further acts in favor of reducing emissions and air pollution [MN19].

A challenge associated with the wide-spread adoption of electric vehicles (EVs) is managing the simultaneous charging demand of large populations of vehicles that might determine violations of constraints in distribution networks and require grid reinforcements [ZMH<sup>+</sup>17, LHS19]. The notion of *smart charging* has been identified as a possible solution to this problem. Smart charging refers to strategies to coordinate the charging process of EVs so as to reduce the peak demand and its effects, see e.g. [IRE19]. EV smart charging to respect grid voltage levels and avoiding power congestions was addressed, e.g., in [SB10, SB11], where authors propose an iterative procedure to determine progressively tighter power constraints for the EVs when grid congestions appear, accounting also for electricity wholesale market prices. The work in [HYLØ13] proposes a market framework for EVs fleet operators, distribution system operators (DSOs), and load balance responsible for satisfying the charging demand of EVs, respecting grid constraints, and providing regulating power to the system. The work in [KSKM17] proposes a centralized scheduler based on a non-convex optimal power flow problem. The work in [MHX<sup>+</sup>20] proposes grid-aware transactive energy management for an EVs fleet. Methods such as rule-based approaches [RS18], heuristics [AAGG14], and central optimization methods [Ric11, SHTT18] have been tested to achieve the goal of effective EV charging management.

In the meanwhile, technologies for autonomous driving are evolving. Autonomous driving is anticipated to disrupt the way we intend transportation and mobility, with implications ranging from ownership

schemes for vehicles (favoring car-sharing and ride-hailing options) to mobility demand (that might increase, fostered by more accessible transportation), see ,e.g., [fCoD15]. When considering large-scale integration of EVs in power grids, autonomous driving will allow vehicles to independently select the most suitable charging locations (e.g., one near a renewable power plant or energy storage facility), offering a new lever to avoid grid congestions. If future mobility is autonomous, grid reinforcements and technological developments planned today for non-autonomous EVs might become obsolete.

The problem of integrating autonomous electric vehicles (AEVs) in distribution grids is not yet entirely explored in the existing literature. The charge scheduling problem for AEVs was addressed in [IMT19] considering the minimization of the waiting times and the electricity costs, neglecting, however, power grid constraints. The problem of planning the charging infrastructure of AEVs was addressed in [LC20, ZSL<sup>+</sup>20] considering mobility patterns of ride-hailing, without considering distribution grids' constraints.

The research summarized in this report refers to integrating the notion of autonomous driving into the smart charging problem of EVs. Two methods that tackle smart charging for AEVs are described. The first is a model-based, centralized approach that assumes perfect knowledge of the distribution grid and demand profiles at grid nodes. However, the absence of perfect information or high uncertainty of the environment might make model-based approaches unpractical or unfeasible (e.g., [AGBB21]). Thus, a second method described in this report considers incomplete or partial knowledge of this information and is based on reinforcement learning.

## 2 A model-based approach using an augmented optimal power flow

This section describes a smart charging scheme for non-autonomous EVs based on an optimal power flow (OPF) applied to a low-voltage distribution grid. In particular, the charging profiles of all the EVs are found by solving an OPF to attain minimum recharging times (to maximize drivers' comfort implicitly) while respecting the operational constraints of the distribution grid. Then, this formulation is modified to accommodate the AEVs. In order to do so, the charging locations of the AEVs, which for the EV case are fixed and depend on where drivers have parked the vehicles, become decision variables of the problem to reflect the fact that AEVs are capable of driving autonomously and picking a charging location that is conducive to improve grid performance.

### 2.1 Problem formulation

Let  $P_{tv}^{(EV)}$  be the charging demand of vehicle  $v$  at time  $t$ . As we consider smart charging, we say that  $P_{tv}^{(EV)}$  is non-negative. The SOC of vehicle  $v = 1, \dots, V$  is:

$$\begin{aligned} \text{SOC}_{tv} \left( P_{tv}^{(EV)} \right) &= \text{SOC}_{t-1v} \left( P_{(t-1)v}^{(EV)} \right) + \eta \frac{1}{E_v} P_{tv}^{(EV)} T_s, \\ &\text{with } P_{tv} \geq 0 \end{aligned} \quad (1)$$

where  $\eta$  is the (constant) charging efficiency,  $E_v$  is the battery energy capacity, and  $T_s$  the sample time in hours. At this stage V2G is not considered, although the the SOC evolution in (1) can be modified to model that

Let  $n = 1, \dots, N$  denote the grid node index. We encode the EVs' charging locations with  $N \times V$  binary variables  $b_{nv}$  that are 1's if vehicle  $v$  charges at node  $n$ , 0's otherwise.

For non-autonomous EVs,  $b_{nv}$ 's are set a-priori (being the charging locations matching with the parking sites) and are an input to the problem. For autonomous EVs, they are free variables of the problem. This is the key difference between the formulation for EVs and AEVs.

The active and reactive power demand at each node of the grid is modelled as the sum of the net demand (i.e., conventional power demand minus distributed stochastic generation, if available) and the aggregated charging demand of all the EVs connected to that node. For time  $t$  and node  $n$ , it is:

$$P_{tn}(P_{t1}, \dots, P_{tV}, b_{n1}, \dots, b_{nV}) = P_{tn}^{(\text{net})} + \sum_{v=1}^V b_{nv} P_{tv}^{(\text{EV})}. \quad (2)$$

The net demand  $P_{tn}^{(\text{net})}$  is an input of the problem and is, for example, from point predictions.

The binary variable  $b_{nv}$  associates the charging demand of vehicle  $v$  to node  $b$  when active. For convenience in the following formulation, we collect all the variables of (2) in the vectors:

$$\mathbf{P}_t^{(\text{EV})} = [P_{t1}^{(\text{EV})}, \dots, P_{tV}^{(\text{EV})}] \quad (3)$$

$$\mathbf{b}_n = [b_{n1}, \dots, b_{nV}]. \quad (4)$$

We assume voltage-independent power demand. As far as the reactive power is concerned, the reactive power of the net demand is also from point predictions (i.e., derived from the active power demand by assuming a certain power factor), whereas the one of EVs is zero as we consider that EV chargers operate at a unitary power factor. The inclusion of the reactive power as a control variable, that might have an impact on voltage profiles in lines with a nonnegligible reactance of the longitudinal parameters (that is not our case study), can be accommodated easily in the formulation and will be considered in future works.

**Grid model** We use load flow equations to model the magnitudes of the nodal voltages and line currents as a function of the active and reactive nodal injections, grid admittance matrix  $Y$  (built from information on the grid topology and on cables parameters), and voltage magnitude at the grid connection point (GCP)  $v_0$ . Let  $v_{tn}$  and  $i_{tl}$  be the voltage magnitude at node  $n$  and current magnitude in line  $l$ , respectively, at time interval  $t$ . We denote the load flow equations by

$$v_{tn}(\mathbf{P}_t^{(\text{EV})}, \mathbf{b}_n) = f_n(P_1(\cdot), \dots, P_N(\cdot), v_0, Y) \quad (5)$$

$$i_{tl}(\mathbf{P}_t^{(\text{EV})}, \mathbf{b}_n) = h_l(P_1(\cdot), \dots, P_N(\cdot), v_0, Y), \quad (6)$$

where we have highlighted the dependencies on the EVs' charging demand and charging locations, which will be the decision variables in the scheduling problem. The complex power absorbed from the (single) GCP is denoted by:

$$S_t(\mathbf{P}_t^{(\text{EV})}, \mathbf{b}_n) = g(P_1(\cdot), \dots, P_N(\cdot), v_0, Y). \quad (7)$$

As known, functions  $f_1, \dots, f_N$ ,  $h_1, \dots, h_L$ , and  $g$  are nonlinear in the power injections lead to nonconvexities and low tractability when integrated in optimal power flows. We resort to sensitivity coefficients using the method described in [CTLBP13] to linearize load flow equations and obtain an approximate solution with a more efficient problem formulation. At this stage, we consider balanced grids, so we carry out a single-phase equivalent load flow.

The problem determines the charging schedule of all vehicles  $v = 1, \dots, V$  over the time horizon  $t = 1, \dots, T$  to achieve a target SOC while respecting all grid constraints. The target state-of-charge of each vehicle  $v$  is denoted by  $\text{SOC}_v^*$ . It should be designed to meet the future driving demand based on driver's input or forecasted based on historical values. In this paper, it is assumed given. By resorting to the models introduced in the previous section, we now formulate the scheduling problems, for traditional EVs first, and for autonomous EVs later.

**Scheduling the charge of non-autonomous EVs** The charging locations for traditional EVs are fully specified by their parking locations. This means that binary variables  $b_{nv}$  are given. These known values are denoted by  $b_{nv}^*$ . The problem consists in determining the charging profiles of all EVs parked at  $b_{nv}^*$  so that the grid constraints are respected and that EVs reach a target state of charge, denoted by  $\text{SOC}_v^*$ , in the least time. This problem reads as:

$$\arg \min_{P_{11}^{(\text{EV})}, \dots, P_{TV}^{(\text{EV})} \in \mathbb{R}_+} \left\{ \sum_{t=1}^T \sum_{v=1}^V \left( \text{SOC}_{tv} \left( P_{tv}^{(\text{EV})} \right) - \text{SOC}_v^* \right)^2 \right\} \quad (8a)$$

subject to EVs' SOC models and power rating limits  $\bar{P}_v^{(\text{EV})}$  of the chargers<sup>1</sup> for  $t = 1, \dots, T$  and  $v = 1, \dots, V$

$$\text{SOC}_{tv} \left( P_{tv}^{(\text{EV})} \right) = \text{SOC}_{t-1v} \left( P_{t-1v}^{(\text{EV})} \right) + \eta \frac{1}{E_v} P_{tv}^{(\text{EV})} T_s \quad (8b)$$

$$0 \leq \text{SOC}_{tv} \leq 1 \quad (8c)$$

$$P_{tv}^{(\text{EV})} \leq \bar{P}_v^{(\text{EV})} \quad (8d)$$

nodal injections model for all  $t, v$ , and  $n = 1, \dots, N$

$$P_{tn} \left( \mathbf{P}_t^{(\text{EV})} \right) = P_{tn}^{(\text{net})} + \sum_{v=1}^V b_{nv}^* P_{tv}^{(\text{EV})} \quad (8e)$$

load flow equations for nodal voltages, lines currents, and power flow at the GCP for all  $t$

$$v_{tn} \left( \mathbf{P}_t^{(\text{EV})} \right) = f_n \left( \mathbf{P}_t^{(\text{EV})}, \mathbf{b}_n^* \right) \quad n = 1, \dots, N \quad (8f)$$

$$i_{tl} \left( \mathbf{P}_t^{(\text{EV})} \right) = h_l \left( \mathbf{P}_t^{(\text{EV})}, \mathbf{b}_n^* \right) \quad l = 1, \dots, L \quad (8g)$$

$$S_t \left( \mathbf{P}_t^{(\text{EV})} \right) = g \left( \mathbf{P}_t^{(\text{EV})}, \mathbf{b}_n^* \right) \quad (8h)$$

which should observe, respectively, statutory voltage levels  $\underline{v}, \bar{v}$ , cables' ampacities  $\bar{i}_l, l = 1, \dots, L$ , and the apparent power rating  $\bar{S}$  at the substation transformer for all  $t$

$$\underline{v} \leq v_{tn}(\cdot) \leq \bar{v} \quad n = 1, \dots, N \quad (8i)$$

$$i_{tl}(\cdot) \leq \bar{i}_l \quad l = 1, \dots, L \quad (8j)$$

$$S_t(\cdot) < \bar{S}. \quad (8k)$$

<sup>1</sup>Reactive power support is not considered at this stage.

As the cost function is convex and all constraints are linear, the optimization problem is convex. We also note that, even if the charging horizon is defined for a fixed time range, one can accommodate arbitrary arrival and departure times by enforcing zero charging power with new linear equality constraints. For example, a predicted departure time at  $t = T - 1$  for vehicle  $v$  can be modeled by adding  $P_{vt}^{(EV)} = 0$  to the constraints.

**Extension to AEVs** Autonomous EVs can pick independently a charging station to accelerate their recharging process and diminishing the impact on the grid. As opposed to the previous problem, the binary variables  $b_{nv}$  are no longer predetermined by the parking locations of the vehicles and are now part of the decision problem, which therefore becomes a mixed integer problem. As the charging locations are now determined by the optimization problem, we need to enforce consistency in the model and ensure that the each vehicle is at one location only. We refer to this requirement as the non-multilocation constraint and it reads as:

$$\sum_{n=1}^N b_{nv} \leq 1. \quad (9)$$

Eq. (2) requires special attention as it features products among decision variables, leading to a complex bi-linear formulation. We use the McCormick's relaxation [McC76] and replace the bi-linear constraint in (2)

$$z_{nvt} = b_{nv} P_{tv}^{(EV)}, \quad (10)$$

where  $b \in \{0, 1\}$  and  $0 \leq P_{nvt}^{(EV)} \leq \bar{P}_v^{(EV)}$ , with the linear inequality constraints

$$z_{nvt} \leq b_{nv} \bar{P}_v^{(EV)} \quad (11a)$$

$$z_{nvt} \leq P_{tv}^{(EV)} \quad (11b)$$

$$z_{nvt} \geq P_{tv}^{(EV)} - \bar{P}_v^{(EV)}(1 - b_{nv}). \quad (11c)$$

As  $b_{nv}$  are binary variables, the relaxation in (11) is exact.

**Additional charge required for autonomous driving** The round-trip drive between the drop-off and charging locations increases the required charging demand to some extent. The additional demand per lag is conservatively estimated by the maximum pairwise distance among grid nodes times the electric energy per unitary distance and is denoted  $E^*$ . As detailed in the following paragraph we will use the same modelling framework as (8), where the binary variables  $b_{nv}$  are let free. To model the requirements of autonomous driving in such a framework, we need to implement the following modeling considerations:

1. For a vehicle to undertake the trip to the charger, its residual SOC at the parking location should be larger than  $E^*$ . If this condition is not met, the vehicle is forced to charge locally by enforcing the respective binary variable  $b_{nv}$  to parking location  $b_{nv}^*$ . Otherwise,  $E^*$  is subtracted from the SOC so that it can be compensated for during the charging process.
2. The target state-of-charge  $SOC_v^*$  is incremented by an amount proportional to  $E^*$  (subject to not incurring in overcharging) so that the energy demand required by the return trip is also compensated for.

3. After, the scheduling problem is completed,  $SOC_v^*$  is curtailed by an amount proportional to  $E^*$  to model the energy spent in the return trip.

These modeling considerations can be implemented with simple pre- and post-optimization heuristics without impacting on the tractability of the problem. More complex partial recharging schemes (e.g., the vehicle achieves a partial charge and drive to a more suitable site for achieving the final charging goal), as well as an improved approximation of the driving distance accounting for road itineraries, will be considered in future works.

**Scheduling the charge of AEVs** The binary variables  $b_{nv}$ , which denote the charging locations of the EVs and were input in the previous formulation, are now decision variables of the problem, leading to a mixed integer program. By using the (exact) relaxation discussed above, we reformulate the bi-linear relationships of (2) into tractable linear constraints. The optimal charging schedules and charging locations  $b_{nv}^o$  are given by

$$\arg \min_{\substack{P_{11}^{(EV)}, \dots, P_{TV}^{(EV)} \in \mathbb{R}_+ \\ b_{11}, \dots, b_{nv} \in \{0, 1\}}} \left\{ \sum_{t=0}^T \sum_{v=1}^V \left( SOC_{tv} \left( P_{tv}^{EV} \right) - SOC_v^* \right)^2 \right\} \quad (12a)$$

subject to the non-multilocation constraint in (9) for all  $n$  and  $v$ , nodal injections and McCormick's exact relaxation for bi-linear constraints for all  $t$ ,  $v$  and  $n$

$$P_{tn} \left( \mathbf{P}_t^{(EV)}, \mathbf{b}_n \right) = P_{tn}^{(net)} + \sum_{v=1}^V z_{nvt} \quad (12b)$$

$$z_{nvt} \leq b_{nv} \bar{P}_v^{(EV)} \quad (12c)$$

$$z_{nvt} \leq P_{tv}^{(EV)} \quad (12d)$$

$$z_{nvt} \geq P_{tv}^{(EV)} - \bar{P}_v^{(EV)} (1 - b_{nv}) \quad (12e)$$

$$P_v^{(EV)} \leq \bar{P}_v^{(EV)}, \quad (12f)$$

and, for all  $t$  and  $v$ , SOC's' evolution as in (8b)-(8c), and load flow's and grid's constraints as in (8f)-(8k) (not reported here for brevity) with the difference that the latter group, depending on the nodal injections, are also a function of the binary decision variables. Additionally, heuristic conditions 1-3 introduced above are implemented by:

$$\text{if } (SOC_{0v} \geq E^*/E_v) : SOC_{0v} = SOC_{0v} - E^*/E_v \quad (12g)$$

$$\text{else} : b_{nv} = b_{nv}^*, \quad \text{for all } v \text{ and } n$$

$$SOC_v^* = \min (SOC_v^* + E^*/E_v, 1), \quad (12h)$$

with a note on (12h)<sup>2</sup>. Once the problem is solved, we subtract from the final SOC the energy required for the return trip from the charging location to the original parking location, if different:

$$\text{if } (b_{nv}^o <> b_{nv}^*) : SOC_{Tv} = SOC_{Tv} - E^*/E_v, \quad (12i)$$

for all  $v$  and  $n$ . Procedures (12g) and (12i) are respectively applied before and after solving the optimization problem based on input and output data and do not alter the problem's properties and tractability.

<sup>2</sup>Eq. (12h) is a (conservative) modeling approximation as the additional demand  $E^*$  should be implemented only when the charging and parking locations are different. The refinement of this modeling aspect will be considered in future works.

## 2.2 Case study

**Distribution Grid** We consider the topology (Fig. 1) and cable characteristics of the CIGRE benchmark system for LV residential grids[CIG09]. The nominal active power demand and power factor are reported in Table 1, and the trajectory of the nodal demand at the node, used to model nodal injections, in Fig. 2. They are all according to specs of the CIGRE benchmark system. The number of vehicles per node, also in Table 1, is set assuming 1.5 vehicles per household, with number of households per node approximated by the nodal nominal demand divided the contractual power for households (e.g., 6 kVA in France). We consider 16 A chargers (i.e., 3.7 kW at nominal voltage). The voltage limits are set to  $1 \text{ pu} \pm 8\%$ , while lines ampacities are according to CIGRE' specs.

Node	Nominal demand (kW)	Power factor	Number of parked EVs	Number of charging AEVs
1	200	0.95	50	54
11	15	0.95	3	6
15	52	0.95	12	5
16	55	0.95	14	15
17	35	0.95	8	10
18	47	0.95	11	6

Table 1: Nominal demand per node and distribution of electric vehicles (the last column refers to results)

**Electric vehicles** We sample the EVs' departing times and SOC at arrival from Weibull (scale 7.67, shape 21.83) and Gaussian (mean 0.49, standard deviation 0.04) distributions, respectively, as shown in Fig. 3. Statistics are estimated from measurements of the *test-an-ev* experiment in Denmark[And13, tes]. For simplicity and to better quantify the impact of autonomous vs. non-autonomous driving, we assume that the arrival time is the same for all vehicles. We adopt the same energy capacity of the EVs used in the experiment, i.e. 16 kWh, so as to retain consistency among collected state-of-charge statistics and driving energy demand.

**Time resolution and scheduling horizon of the optimization problem** Resolution time is 1 hour; based on the min/max values of the arrival times distribution (approximated to the nearest integer hour in-line with the adopted time resolution), we set the scheduling horizon of the problem to be from 16h to 8h of the next day.

**EVs' charging objectives** The target state-of-charge implemented in the cost functions of problems (8) and (12) is  $SOC_v^* = 100\%$  for all vehicles  $v = 1, \dots, V$ . The additional charge  $E^*$  for accomplishing the autonomous drive to the charging station is the the maximum distance among nodes (345 meters, N18-N1 in Fig. 1) times the average consumption per km (0.160 kWh/km from experimental data[Cle14]). It amounts to 55 Wh, i.e. 0.8% SOC.

**Key performance indicators** Evaluating the cost function of the optimization problems offers an immediate interpretation of the performance of the scheduler along the charging horizon. Based on it, the first metric is:

$$\text{Metric 1} = \sum_{t=1}^T \sum_{v=1}^V (SOC_{tv} - SOC_{tv}^*)^2 \quad (13)$$

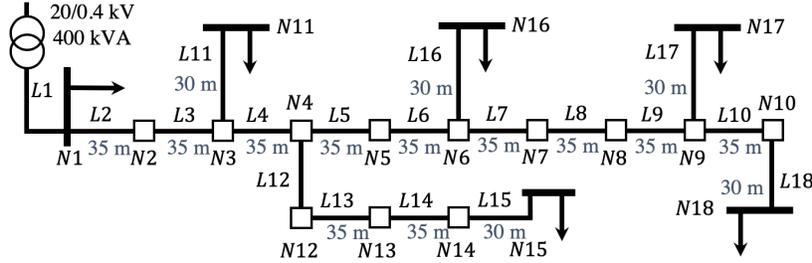


Figure 1: Topology of European LV distribution network benchmark for residential system used for the verification[CIG09].

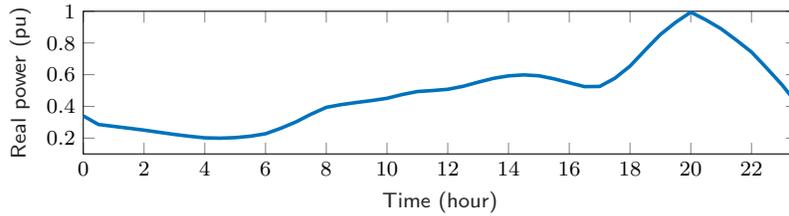


Figure 2: Profile of the net active power demand from CIGRE specifications[CIG09].

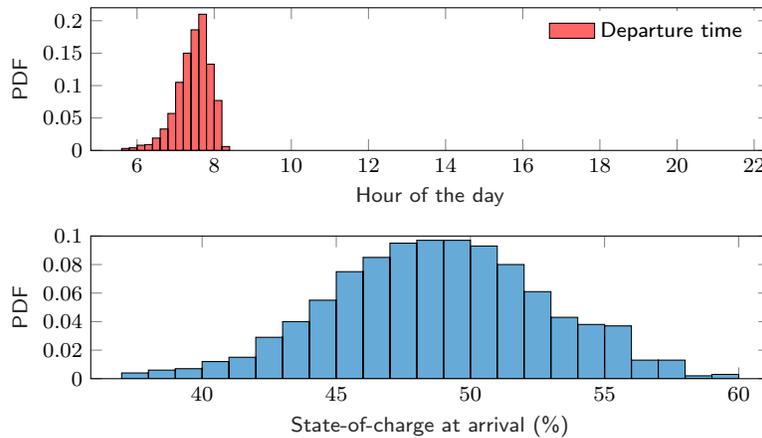


Figure 3: Distributions of the EVs' departure times (upper-panel plot) and of the initial SOC (lower-panel plot) used in the simulations.

The second metric evaluates the performance at the end of the scheduling period  $T$  and measures if target charging objectives have been met:

$$\text{Metric 2} = \sum_{v=1}^V (SOC_{Tv}^* - SOC_{Tv}). \quad (14)$$

Finally, Metric 3 is the time taken by the charging process to reach the target SOC for all vehicles.

## 2.3 Results

The number of AEVs that are charged at the various nodes of the grid is shown in the last column of Table 1. It can be seen that, compared to the case of non-autonomous EVs, 13 AEVs pick a different charging location. As to be expected, AEVs tend to favor charging locations closer to the GCP, that allow connecting higher demand with less impact on voltage levels. For example, nodes 1 and 11 feature additional 4 and 3 charging vehicles than parked vehicles, as opposed to remote nodes 15 and 18 that are with less vehicles.

The upper- and middle-panel plots of Fig. 4 show the charging power and SOC's evolution at different quantiles of the population of EVs, respectively. AEVs feature a higher degree of simultaneity when charging and achieve the target SOC quicker than conventional EVs. Thanks to picking charging locations closer to the GCP, AEVs achieve simultaneous charging while respecting voltage constraints as visible in the middle-panel plot of Fig. 4.

Table 2 shows that autonomous EVs achieve a lower realization of the cost function (Metric 1) and the target SOC in 4 hours, nearly half of the time than conventional vehicles (Metric 3)<sup>3</sup>. Conventional EVs score better in Metric 2 because those autonomous vehicles which change locations for charging cannot achieve 100% SOC as some charge is spent in driving back to the original parking location.

Metric	Conventional	Autonomous
Metric 1	373	300
Metric 2	0%	0.8%
Metric 3	7 hour	4 hour

Table 2: Metrics for conventional vs. autonomous vehicles.

## 3 Model-free stochastic control with deep reinforcement learning

This section presents a method and a use case that demonstrates the application of deep reinforcement learning (DRL) to control the charging power at an electric and autonomous vehicle (AEV) charging node. Based on the analyses, we demonstrate the capability of DRL to learn the optimal charging policy in a highly stochastic and partially observable environment with multiple charging objectives. Moreover, we derive the state vectors based on the readily available observations through standard metering infrastructure in a low-voltage (LV) network and perform batch-wise learning via policy gradient update.

<sup>3</sup>As the distance among nodes is small (i.e., < 350 meters), we neglect the time it takes to drive from the charging location to the parking location

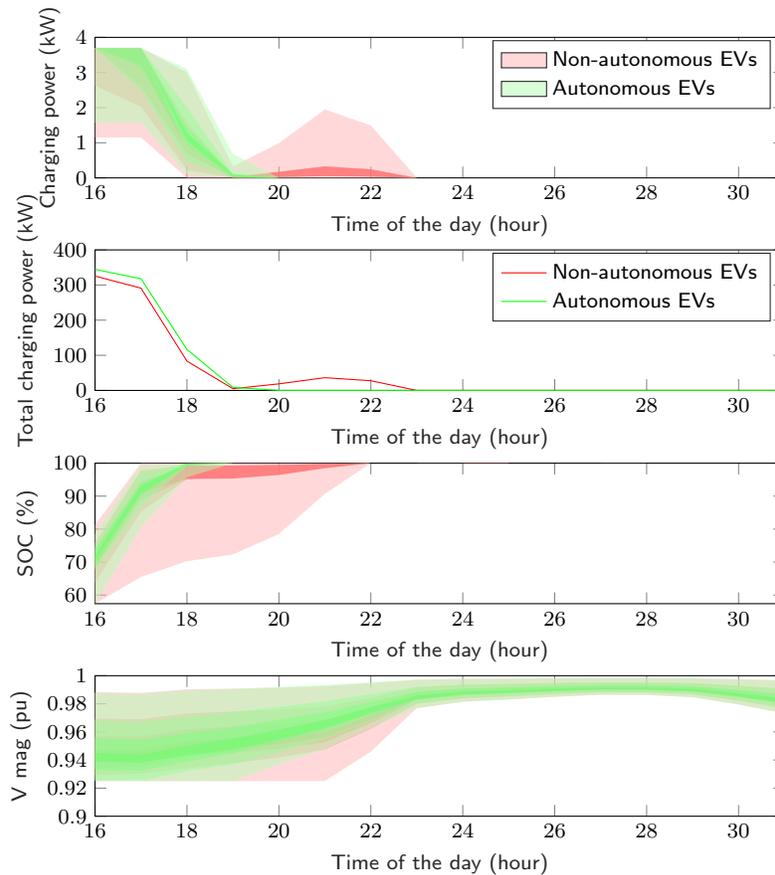


Figure 4: Charging power (first panel), total charging power, and SOC (third panel) of the EVs population, and nodal voltage magnitude (fourth panel) of all grid nodes. In the 1st, 3rd, and 4th plots, the lighter color-shade refers to the 0 and 1 quantiles, the thicker to the 0.43 and 0.57 ones, whereas the two primary colors red and green refer to conventional and autonomous EVs, respectively.

### 3.1 A motivating example for partial observability in smart-grid

To demonstrate the effect of partial information on smart-grid control, we present the following oversimplified example. Figure 5 shows a simple, radial network with two loads and one controllable generator. We attempt to control the generator such that the generator node voltage is above the statutory limit of 0.95 p.u. using only the measurements at the generator node.

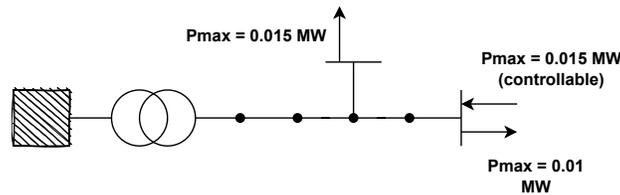


Figure 5: The motivating example for showing the effect of partial information on smart-grid control.

The mathematical formulation of this simple problem is similar to [GSP22] except we estimate the voltage sensitivity coefficient based on the local voltage and power measurements. We estimate the voltage coefficient using Ridge regression with a  $R^2$  score of 0.86. Figure 6 depicts the outcome of the control scheme after testing it for 1000 time steps using stochastic, normalized load profiles. Despite having a strict constraint on the voltage limit, the controller cannot generate an accurate control signal. The root cause of the problem is related to the inaccurate voltage sensitivity estimation due to partial information. Consequently, the voltage sensitivity estimation ignores the coupling effects of the LV grid and is biased towards the local observations used to calculate the voltage sensitivity coefficient.

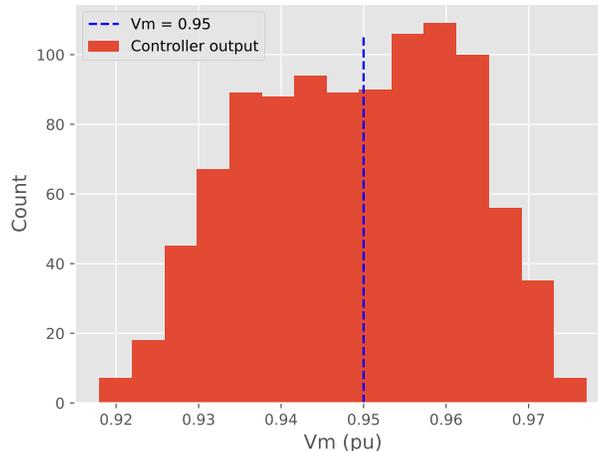


Figure 6: The motivating example for showing the effect of uncertainty and partial information on smart-grid control.

While the advantage of having complete information for optimal control decisions is clear, we often need to build our solutions based on partial data due to the cost of installing widespread advanced metering infrastructure. Scientists are conducting extensive research on optimal sensor placement that looks at the trade-offs between the information gain and investment cost of additional metering infrastructure. Nevertheless, the constant expansion of renewable generators, smart energy devices, and

controllers have radically increased the non-stationarity of the smart-grid environment, wherein optimal trade-off scenarios can constantly change over time. Therefore, there is a need for future smart-grid control algorithms to be explicitly more robust towards the partial information and non-stationary conditions of the smart-grid environment. A recent review by Tàczi et al. [TSVH21] also finds similar gaps in current smart-grid research.

As a result, there has been an increasing interest in more flexible data-driven approaches to smart-grid applications such as managing AEV charging. Data-driven strategies can be used without assumptions regarding the underlying model, and they can represent the inherent stochasticities in the environment. Moreover, data-driven methods enable us to learn stochastic strategies that outperform deterministic strategies over long time horizons, even in adversarial settings [WBH<sup>+</sup>16].

### 3.2 Theoretical background

Our objective is to regulate the AEV charging power to minimize the charging time and voltage limit violations at the charging node. The power flow equations describe the relationship between power and voltage in an electrical distribution network. For simplicity, we do not consider reactive power control in our use case. However, it is important to note that the European LV grid benchmark has R/X ratios of 0.7 - 11.0 [AAAT18], which are relatively high, and at high R/X ratios, active power has the most significant influence on voltage [BP08]

The mathematical form of the objective function is given by Equation 15. In Equation 15,  $P_{max}$  is the maximum charging load (maximum charging power of a charging point  $\times$  number of charging points at the node),  $\alpha^t = P_c^t/P_{max}$  is the ratio between the charging load at time  $t$  and  $P_{max}$ ,  $\mathcal{N}$  is the set of nodes in the LV grid,  $V_m$  is the voltage magnitude at the charging node, and  $V_{lb}$  is the statutory voltage limit.

$$\begin{aligned}
& \max_{\alpha^t} \quad \mathbb{E}_{t \in \mathcal{T}} \left( \mathbb{1}^{|V_m^t - V_{lb}| \leq \zeta} + \mathbb{1}^{V_m^t > V_{lb} + \zeta} \alpha^t \right) \\
& \text{s.t.} \quad P_i^t = |V_i^t| \sum_{j \in \mathcal{N}} |V_j^t| (G_{ij} \cos \delta_{ij}^t + B_{ij} \sin \delta_{ij}^t) \\
& \quad \quad Q_i^t = |V_i^t| \sum_{j \in \mathcal{N}} |V_j^t| (G_{ij} \sin \delta_{ij}^t - B_{ij} \cos \delta_{ij}^t) \\
& \quad \quad P_c^t = \alpha^t P_{max}
\end{aligned} \tag{15}$$

Equation 15 is a concise way to combine both the charging power and voltage objectives. The statutory voltage limit is imposed as a soft constraint with a small allowable margin of error of  $\zeta$ .

To solve the optimization problem with DRL, we need to define the states, actions, and reward function of the reinforcement learning (RL) agent. Moreover, we employ the stochastic policy gradient approach that enables us to create an RL agent that learns the optimal stochastic policy directly from observations. There are a variety of policy gradient algorithms published in the literature. The algorithm used in our case studies is called proximal policy optimization (PPO), which was first published in 2017 [SWD<sup>+</sup>17]. The main advantages of PPO are its simplicity and general applicability. PPO belongs to the class of algorithms known as actor-critic, where two function approximators (the critic and the actor) work in tandem to learn the value function and the optimal policy. The actor-network is a parameterized representation of the agent's current policy  $\pi$ . At each iteration, the agent takes action based on the state of the environment, and its current policy, i.e.,  $a^t = \pi(s^t)$ . The critic evaluates the value of the action at the given state and updates the value function's parameters using a temporal difference

update. Finally, the actor updates the policy in the policy-gradient direction, where the gradient is calculated using the critic's value estimate.

**States** We define two state vectors, one for the critic and another for the actor. The state vector of the actor is a local subset of the state vector available to the critic, which imposes the partial observability condition.

Critic's state, denoted by  $S_c$ , is a discrete-transformed vector of voltage magnitudes (in p.u.) at each load and generator connected bus. Given the number of load-connected buses is  $m$ ,  $n$  is the number of generator-connected buses, and  $l$  is the number of bins, the critic state at time  $t$  is a vector of the shape  $(1, m + n, l)$ . The actor's observability is limited to the charging node itself and the  $b$  nearest load or generator buses from the charging node. Therefore, the actor state is a vector of the shape  $(1, b, l)$ <sup>4</sup>.

**Actions** The agent policy yields an action at each time-step  $t$  that regulates the charging power. Therefore, we define the action of the stochastic charging agent as  $\alpha^t = \pi(s^t)$ . Clearly,  $\alpha^t$  is a real value in the range  $[0, 1]$  that can be represented as a random realization of a beta policy, i.e.,  $\alpha^t \sim \text{Beta}(a, b)$ . In other words, we can write the optimal stochastic policy  $\pi^* = \text{Beta}(a^*, b^*)$  where  $a^*, b^*$  are the optimal parameter values of the beta policy.

**Reward function** Following Equation 15, we define the reward function as;

$$R(s^t, a^t) = \mathbb{E} \left( \mathbb{1}^{|V_m - V_{ib}| \leq \zeta} + \alpha^t (\mathbb{1}^{V_m > V_{ib} + \zeta}) \right) \quad (16)$$

It is important to note that when the environment consists of multiple charging stations,  $\alpha^t$  becomes a vector of actions. The reward is also a vector of rewards corresponding to each charging agent; therefore, we take the expected reward as the training signal for the critic network. Moreover, the role of the critic is only limited to enabling the training process of the actor networks. During the policy roll-out, we disconnect the critic and let the decentralized actors to take optimal actions based on their learned policy.

We invite the reader to refer to [HN22] for more comprehensive details regarding the policy gradient formulation of the optimization problem.

So far, we have designed a mathematical formulation that enables us to optimally control the total charging power at a node minimizing expected voltage violations and charging time. The assignment problem that we discuss now answers the question of the equitable allocation of the total charging power between the multiple vehicles that require charging simultaneously. We define equity as minimizing the sum of instantaneously evaluated charging times for all vehicles. This definition allows us to prioritize more depleted AEVs and charge them faster. Consequently, we expect more AEVs to be available for users, leading to better mobility services. The non-linear optimal power assignment problem can be written as in Equation 17, where  $K'$  is the set of active charging points. Furthermore,  $\alpha^{k',t}$  is the charge rate of the charging point  $k'$  at time  $t$ , and it is a real value in the range  $[\epsilon, 1]$ . The lower-bound  $\epsilon$  is a very small real value introduced for numerical stability.

---

<sup>4</sup>In our case study  $b = 3$ .

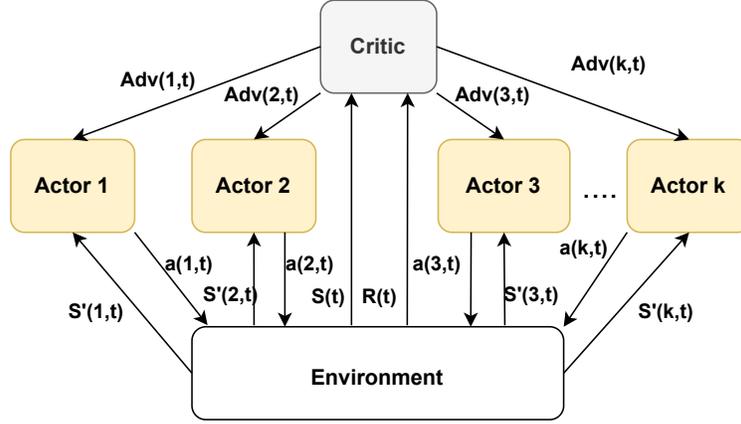


Figure 7: The actor-critic architecture implemented by us is one with a centralized critic and decentralized actors where the critic has full observability of the environment and each actor receives only a subset of local information. The role of the critic is to enable the learning process of actors.

$$\begin{aligned}
 \min_{\alpha^{k',t}} \quad & \frac{1 - SOC^{k',t}}{\alpha^{k',t} + \epsilon} \\
 \text{s.t.} \quad & 0 \leq \alpha^t P_{max} - \sum_{k' \in \mathcal{K}'} \alpha^{k',t} P_{max}^{k'} \\
 & \alpha^{k',t} \leq \epsilon \quad \text{if } SOC^{k',t} = 1 \\
 & \epsilon \leq \alpha^{k',t} \leq 1
 \end{aligned} \tag{17}$$

Figure 8 is a pictorial depiction of the problem formulation that shows the solution steps at one time instance.

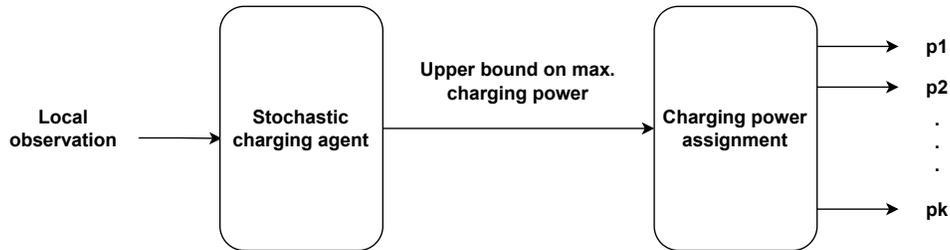


Figure 8: Flow diagram that shows the interconnection of outer and inner optimization problems at a given time step. We run this process in iteration for each time step of the simulation.

### 3.3 Case study 1

The first case study consists of 216 trips within the Swiss municipality Lugano within a day. The travel data is synthetically generated using MATsim (<http://www.matsim.org>), an agent-based micro-simulation framework for mobility systems simulations [HNA16]. The road network extracted from OpenStreetMaps as a graph contains all roads and links in Lugano with the importance level either

residential or higher. The metadata includes distance and maximum travel speed for each edge of the graph. The resulting network has 1122 nodes and 3602 edges.

To simulate the power system impacts, we use a modified CIGRE LV benchmark grid (Figure 9) with representative residential load profiles. The residential load profiles are obtained by simulating typical household appliances and devices (heat pumps and boilers, rooftop PV generation, and non-dispatchable demand). The charging node is a meta-node with 11 charging points with a maximum charging power of 11kW (L19 in the CIGRE benchmark grid).

A discrete-time simulation environment with one minute time resolution based on SimPy [Mat08] is developed to simulate the fleet of shared AEVs servicing the travel requests. The fleet consists of 11 AEVs, and they are randomly located at the start of the simulation. A python generator pops a travel request when the environment time reaches the start time of a trip. A free AEV can accept that request and initiate a series of processes to service the request by 1) routing to the pickup location, 2) picking up the customer, and 3) routing to the destination. En route, an AEV can decide to charge the batteries if it senses a chance of battery depletion. Similarly, an AEV can leave the charging station during the charging process when it senses sufficient SOC to serve an incoming travel request.

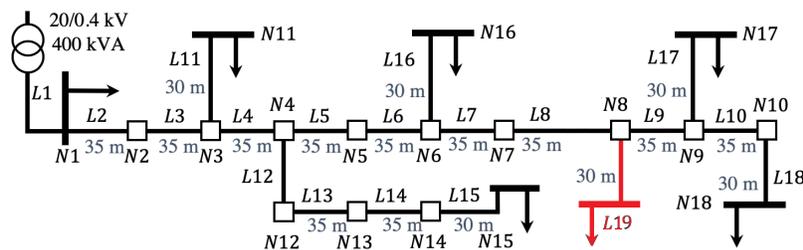


Figure 9: Modified CIGRE LV grid used in the case study 1. L19 represent the charging station with 11 charging points. Figure adapted from [CIG09].

The training dataset consists of 20 days of residential load profiles covering all four seasons of the year. We added a small Gaussian noise to each residential load profile during model training to enable the stochastic charging agent to learn from similar but not identical observations at each iteration. The customer travel demand profiles are identical in each day. The validation dataset consists of 10 days of residential load profiles (without added Gaussian noise), and the customer demand profiles are identical to the one in the training data. The model training is performed in batches of 64 randomly sampled observations from a replay buffer. For additional details regarding the model hyper-parameters, training, and testing, please refer to [HN22].

### 3.4 Results - case study 1

The stochastic charging agent predicts a charging power upper bound with a mean of approximately 68% of the maximum charging load of the station (Figure 10). In the following, we present the performance of the stochastic charging agent in comparison to two scenarios.

- Scenario in which vehicles are always charged at the maximum power,
- Scenario in which the charging power is regulated as a simple function of the node voltage as given below.  $\zeta$  is set to 0.01 in our case study.

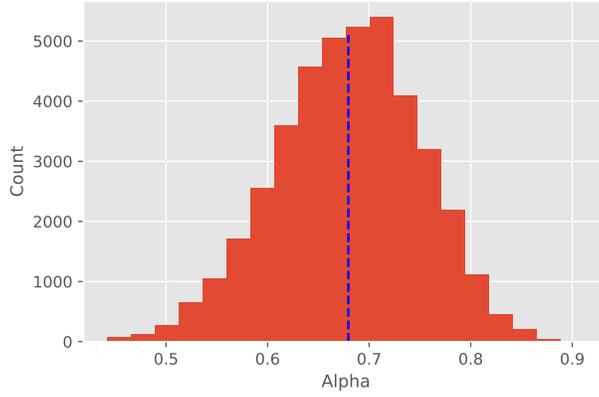


Figure 10: The distribution of the stochastic charging agent's predictions over the validation period.

$$\alpha^t = \begin{cases} \frac{V_m - V_{lb} - \zeta}{V_{max} - V_{lb} - \zeta} & V_{lb} - \zeta \leq V_m \\ 0.5 & otherwise \end{cases}$$

The peak shaving effect takes place only at specific times of the day when the charging power demand exceeds the upper bound forecast of the stochastic charging agent, as shown in Figure 11a. We also observe, in comparison to the benchmark strategy, that the stochastic charging agent enforces higher charging rates when possible (Figure 11b). The voltage impact of peak-shaving is depicted in Figure 12. Over the 10-day validation period, the stochastic control strategy results in 17 instances of voltage dead-band violations (0.12% of the total observed time steps), whereas the benchmark strategy results in zero violations. However, the proposed strategy provides a 7.4% extra charging rate during the same period, on average. Furthermore, between the peak charging times (time steps 400 to 1000 of each day), the proposed strategy provides an additional 39.07% average charging rate compared to the benchmark strategy.

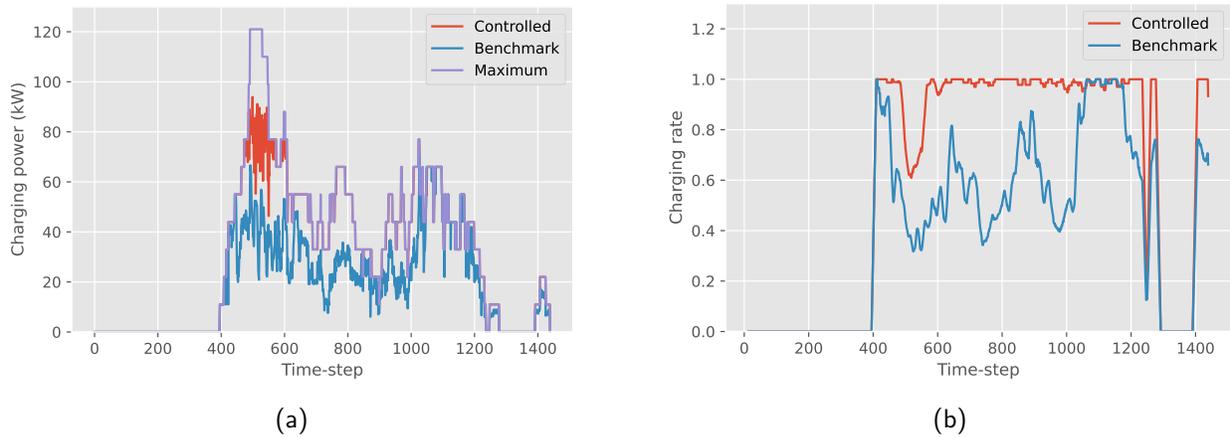


Figure 11: (a) The peak shaving effect of the stochastic charging agent compared with the benchmark strategy and no-control strategy, (b) The charge rate forecasted by the stochastic agent and with benchmark strategy for one validation day.

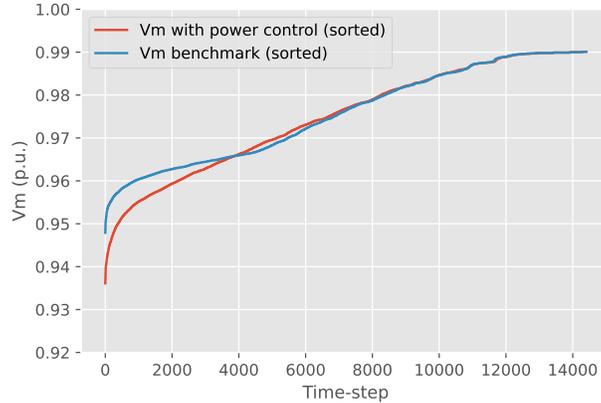


Figure 12: The voltage magnitudes at the charging node for the 10-day period using the proposed and benchmark strategies, sorted in the ascending order

Figure 13a is a graphical depiction of how the state of charge (SOC), charging rates, and  $\alpha^t$  are related to each other. Firstly, we observe that the charging rates increase when the SOC is lower, which is the expected behavior of the inner optimization. However, the sensitivity of this relationship is governed by  $\alpha^t$ . If the constraint is strict (low  $\alpha^t$ ), the charging rate becomes more sensitive to the changes in SOC. Conversely, if the charging power constraint is lenient, the sensitivity of the charging rate to SOC gets lower.

The charging trajectories (profiles) describe the change of SOC of a vehicle over time (Figure 13b). Due to the negative dependency of the charging rates on SOC, the charging profiles of the AEVs are, by default, non-linear. Charging trajectories can progress linearly only when the total charging power requirement is less than the constraint set by the stochastic charging agent. The non-linearity of the charging profiles exacerbate when the charging power constraint is more stringent. As a result, as the SOC of a vehicle increases beyond a certain threshold, it may become unproductive for an AEV to remain connected to the charging point, given the diminishing charging rates. As a result, this behavior provides an additional degree of freedom for intelligent decision-making and optimization. For example, we can argue that in a sharing economy, it is much better to have two vehicles at 70% SOC levels than to have one vehicle fully charged and the other one at, say, 40%. The additional degree of freedom encourages faster turnover of vehicles and can improve the use of limited charging resources. While we do not address this question in the current article, we would like to present it to the research community as a promising area to investigate.

### 3.5 Case study 2

Our second case study is an environment with 740 trips within one day in the Swiss canton Ticino as a proof of concept for a complex environment. The corresponding graph object of the road network in Ticino has 10286 nodes and 33519 edges. The data sources are the same as in case study 1.

The LV grid is a benchmark grid (code: 1-LV-semiurb4–2-no\_sw) provided in the SimBench dataset [MSD<sup>+</sup>20]. The environment has five charging stations, each with 11 charging points, and they are connected to the buses 6, 12, 18, 24, and 38 of the benchmark grid. The maximum number of AEVs in the simulation environment is 55. Their locations (with respect to the road network) are determined semi-randomly. To ensure charging station locations are not geographically biased, we manually create

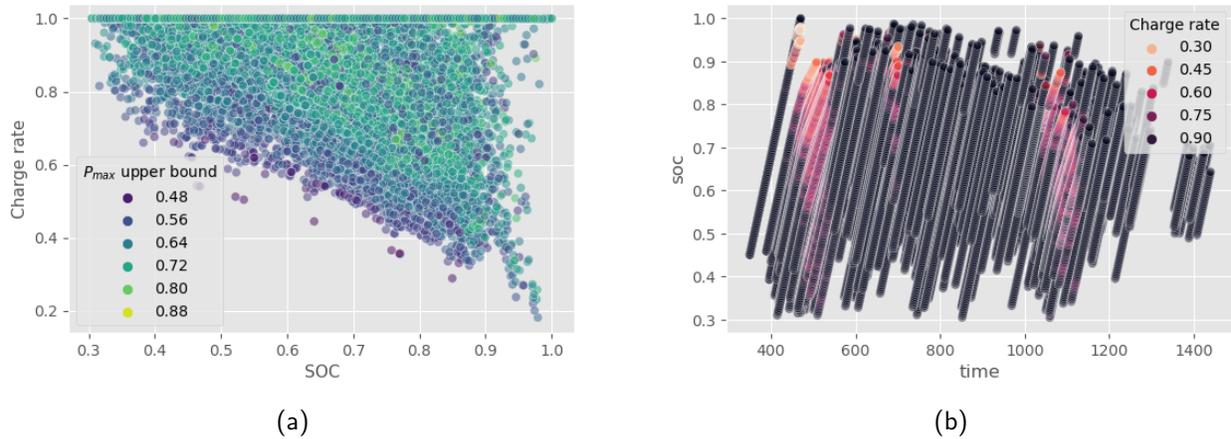


Figure 13: (a) Relationship between the SOC and charging rates, (b) Charging trajectories of the AEVs

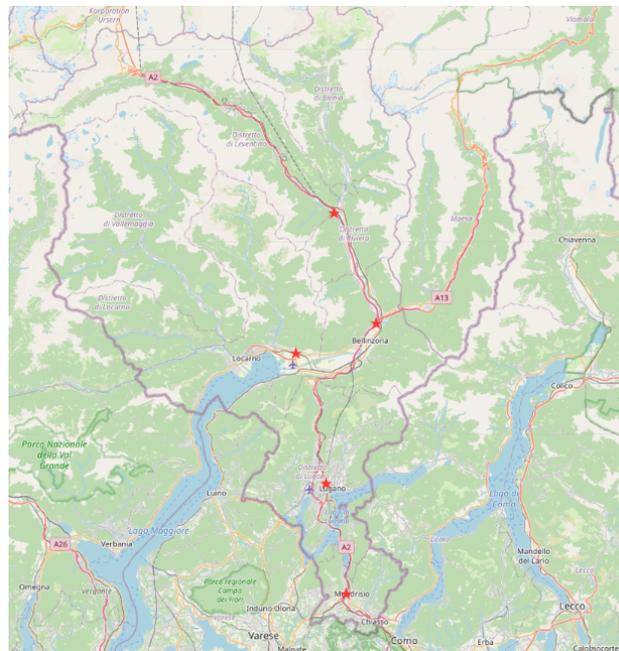


Figure 14: Approximate locations of the charging stations in case study 2.

five regional clusters of traffic nodes and randomly choose one node from each regional cluster to be a charging nodes. Approximate locations of the charging stations are shown in Figure 14. The residential load profiles are generated using the same strategy described earlier, while charging station and charging point properties remain identical to case study 1.

### 3.6 Results - case study 2

First of all, we observe the voltage regulation effect of the stochastic charging agent in Figure 15. The results show that there are 20 and 29 instances of voltage violations (0.13% and 0.20% of the total

observed time steps) in the buses 24 and 6, respectively. These two buses are located at the far end of the longest feeder in the SimBench LV benchmark grid [MSD+20] and consequently become more prone to voltage violations.

The observation also highlights one of the potential weaknesses of stochastic controllers, i.e., they cannot strictly impose hard constraints that are particularly important for safety-critical applications. The typical ways in which the constraints are handled in RL problems is to either add a penalty function that discourages the gradient updates in the undesired direction or define threshold probability for the constraint violation [CQT+22]. Although on average, both strategies minimize the chance of an agent taking an action that leads to an undesirable state, they do not guarantee complete satisfaction of the constraint at all times. On the other hand, it is also possible for an agent to take conservative actions to ensure that the constraints are satisfied at all times, which may lead to the overall conservativeness of the agent [CQT+22].

Our case studies, being not safety-critical, use the soft constraint approach with a penalty function. However, we observed improved performance of the stochastic charging agent in reducing the number of constraint violations under the following conditions.

- Longer training periods that, even intuitively, enables the agent to learn more about its environment,
- Assisted exploration, by which we add a small Gaussian noise around the residential load profiles and synthetically allow the RL agent to observe states that are otherwise not revealed to it by nature,
- Exploration of the action space, by which we add an entropy term to the actor’s loss function (entropy regularization) and force the RL agent to try different actions given the similar states.

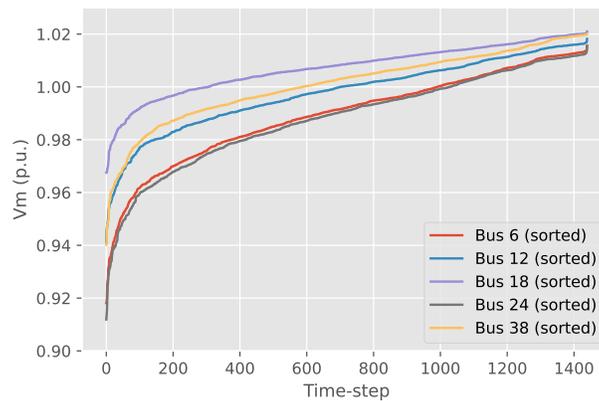


Figure 15: The voltage magnitudes at the charging nodes for the 10-day period.

We also observe a significant increase in total travel distance under the benchmark control scheme (Figure 16a). As we recall from case study 1, the benchmark controller does not optimize the charging power, and particularly during peak charging times, it delivers less charging power than the stochastic controller. As a result, under the benchmark control scheme, AEVs are required to charge more often (Figure 16b) and divert more times to find charging locations, resulting in sub-optimal routing. It is important to note that this increase in unwanted travel distance and the number of charging events

depends heavily on the number of charging stations in the environment and their placement. However, it is clear from the observations that the optimality of the charging control scheme has much broader implications concerning energy security, economic efficiency, and (as discussed in case study 1) social equity.

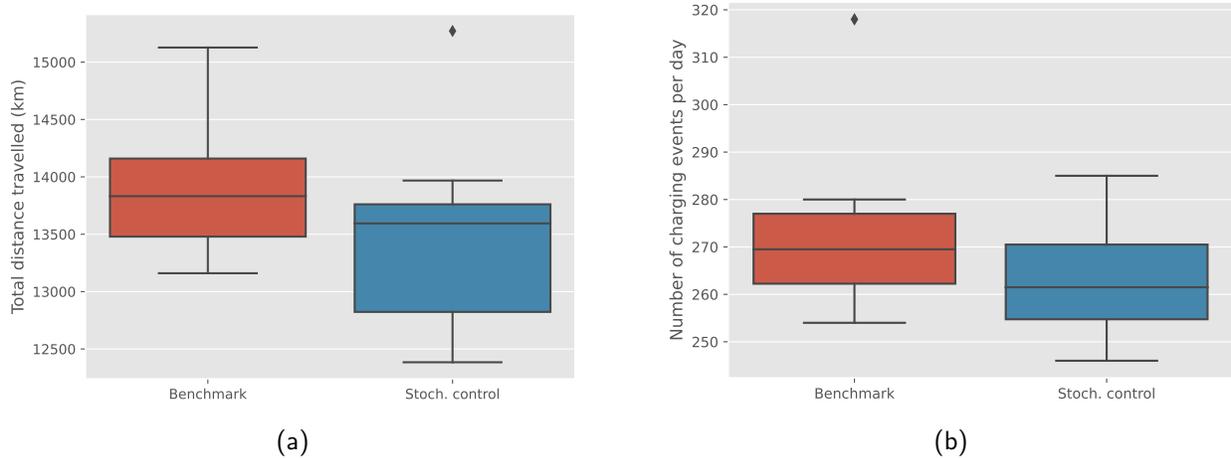


Figure 16: (a) Comparison of total travel distance in a day and (b) Comparison of the number of charging events in a day with simulation results of the 10-day validation period.

## List of publications

Peer-reviewed conference papers:

- Sossan, Fabrizio, Biswarup Mukherjee, and Zechun Hu. "Impact of the charging demand of electric vehicles on distribution grids: a comparison between autonomous and non-autonomous driving." 2020 Fifteenth International Conference on Ecological Vehicles and Renewable Energies (EVER). IEEE, 2020.
- Mukherjee, Biswarup, Georges Kariniotakis, and Fabrizio Sossan. "Smart Charging, Vehicle-to-Grid, and Reactive Power Support from Electric Vehicles in Distribution Grids: A Performance Comparison." 2021 IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe). IEEE, 2021.
- Mukherjee, Biswarup, Georges Kariniotakis, and Fabrizio Sossan. "Scheduling the Charge of Electric Vehicles Including Reactive Power Support: Application to a Medium-Voltage Grid." (2021): 1534-1538.
- Heendeniya, Charitha Buddhika and Nespoli, Lorenzo. "A stochastic deep reinforcement learning agent for grid-friendly electric vehicle charging management." In 2022 DACH+ Energy Informatics conference, 2022. (Accepted)

## References

- [AAAT18] Melike Selcen Ayaz, Rasoul Azizipanah-Abarghooee, and Vladimir Terzija. European LV microgrid benchmark network: Development and frequency response analysis. *2018 IEEE International Energy Conference, ENERGYCON 2018*, pages 1–6, 2018.
- [AAGG14] Monica Alonso, Hortensia Amaris, Jean Gardy Germain, and Juan Manuel Galan. Optimal charging scheduling of electric vehicles in smart grids by heuristic algorithms. *Energies*, 7(4):2449–2475, 2014.
- [AGBB21] Heba M. Abdullah, Adel Gastli, and Lazhar Ben-Brahim. Reinforcement Learning Based EV Charging Management Systems-A Review. *IEEE Access*, 9:41506–41531, 2021.
- [And13] Peter Bach Andersen. *Intelligent electric vehicle integration-domain interfaces and supporting informatics*. PhD thesis, Technical University of Denmark (DTU), 2013.
- [BP08] B. Blažič and I. Papič. Voltage profile support in distribution networks - Influence of the network R/X ratio. *2008 13th International Power Electronics and Motion Control Conference, EPE-PEMC 2008*, pages 2510–2515, 2008.
- [CBER21] Francesca Cellina, Albedo Bettini, D Eva, and Roman Rudel. Literature review regarding future mobility scenarios. Technical report, SUPSI, Mendrisio, 2021.
- [CIG09] CIGRE' Task Force C6.04.02. Benchmark systems for network integration of renewable and distributed energy resources. Technical report, CIGRE International Council on large electric systems, July 2009.
- [Cle14] Clever A/S. Test-en-elbil slutrapport. Technical report, Clever A/S, July 2014.
- [CQT<sup>+</sup>22] Xin Chen, Guannan Qu, Yujie Tang, Steven Low, and Na Li. Reinforcement Learning for Selective Key Applications in Power Systems: Recent Advances and Future Challenges. *IEEE Transactions on Smart Grid*, 3053(c):1–25, 2022.
- [CTLBP13] K. Christakou, D.-C. Tomozei, J.-Y. Le Boudec, and M. Paolone. Gecn: Primary voltage control for active distribution networks via real-time demand-response. *IEEE Transactions on Smart Grid*, 4(2):622–631, 2013.
- [fECOD15] Organisation for Economic Co-operation and Development. *Urban mobility system upgrade: How shared self-driving cars could change city traffic*. OECD Publishing, 2015.
- [GSP22] Rahul Gupta, Fabrizio Sossan, and Mario Paolone. Model-less robust voltage control in active distribution networks using sensitivity coefficients estimated from measurements. *arXiv preprint arXiv:2201.04192*, 2022.
- [HN22] Charitha Buddhika Heendeniya and Lorenzo Nespoli. A stochastic deep reinforcement learning agent for grid-friendly electric vehicle charging management. In *2022 DACH+ Energy Informatics conference, 2022*. (Accepted).
- [HNA16] Andreas Horni, Kai Nagel, and Kay W Axhausen. *Introducing MATsim*. Ubiquity Press, London, 2016.

- [HYLØ13] Junjie Hu, Shi You, Morten Lind, and Jacob Østergaard. Coordinated charging of electric vehicles for congestion prevention in the distribution grid. *IEEE Transactions on Smart Grid*, 5(2):703–711, 2013.
- [IMT19] Riccardo Iacobucci, Benjamin McLellan, and Tetsuo Tezuka. Optimization of shared autonomous electric vehicles operations with charge scheduling and vehicle-to-grid. *Transportation Research Part C: Emerging Technologies*, 100:34–52, 2019.
- [IRE19] IRENA. Innovation outlook: Smart charging for electric vehicles. Technical Report MSU-CSE-06-2, International Renewable Energy Agency, Abu Dhabi, 2019.
- [KSKM17] Katarina Knezović, Alireza Soroudi, Andrew Keane, and Mattia Marinelli. Robust multi-objective pq scheduling for electric vehicles in flexible unbalanced distribution grids. *IET Generation, Transmission & Distribution*, 11(16):4031–4040, 2017.
- [LC20] Mustafa Lokhandwala and Hua Cai. Siting charging stations for electric vehicle adoption in shared autonomous fleets. *Transportation Research Part D: Transport and Environment*, 80:102231, 2020.
- [LHS19] Z. Lin, Z. Hu, and Y. Song. Optimal planning of paired pev charging stations and esss in distribution networks considering flexibility of charging loads. In *2019 IEEE Power Energy Society General Meeting (PESGM)*, pages 1–5, Aug 2019.
- [Mat08] Norm Matloff. Introduction to discrete-event simulation and the simpy language. *Davis, CA. Dept of Computer Science. University of California at Davis. Retrieved on August, 2(2009):1–33*, 2008.
- [McC76] Garth P McCormick. Computability of global solutions to factorable nonconvex programs: Part i—convex underestimating problems. *Mathematical programming*, 10(1):147–175, 1976.
- [MHX<sup>+</sup>20] A. Masood, J. Hu, A. Xin, A. R. Sayed, and G. Yang. Transactive energy for aggregated electric vehicles to reduce system peak load considering network constraints. *IEEE Access*, 8:31519–31529, 2020.
- [MN19] Richard Mounce and John D. Nelson. On the potential for one-way electric vehicle car-sharing in future mobility systems. *Transportation Research Part A: Policy and Practice*, 120:17–30, 2019.
- [MSD<sup>+</sup>20] Steffen Meinecke, Džanan Sarajlić, Simon Ruben Drauz, Annika Klettke, Lars Peter Lauven, Christian Rehtanz, Albert Moser, and Martin Braun. SimBench-A benchmark dataset of electric power systems to compare innovative solutions based on power flow analysis. *Energies*, 13(12), 2020.
- [Ric11] A. Richardson, P.; Flynn, D.; Keane. Optimal Charging of Electric Vehicles in Low Voltage Distribution Systems. *IEEE Transactions on Power Systems*, 27(1):268–279, 2011.
- [RS18] Abdul Rauf and Zainal Salam. A rule-based energy management scheme for uninterrupted electric vehicles charging at constant price using photovoltaic-grid system. *Renewable Energy*, 125:384–400, 2018.

- [SB10] Olle Sundström and Carl Binding. Planning electric-drive vehicle charging under constrained grid conditions. *2010 International Conference on Power System Technology: Technological Innovations Making Power Grid Smarter, POWERCON2010*, pages 0–5, 2010.
- [SB11] Olle Sundstrom and Carl Binding. Flexible charging optimization for electric vehicles considering distribution grid constraints. *IEEE Transactions on Smart Grid*, 3(1):26–37, 2011.
- [SHTT18] Bo Sun, Zhe Huang, Xiaoqi Tan, and Danny H K Tsang. Optimal Scheduling for Electric Vehicle Charging With Discrete Charging Levels in Distribution Grid. *IEEE Transactions on Smart Grid*, 9(2):624–634, 2018.
- [SWD<sup>+</sup>17] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal Policy Optimization Algorithms. pages 1–12, 2017.
- [tes] Test-an-ev project: Electrical vehicle (ev) data. <http://mclabprojects.di.uniroma1.it/smarthgnew/Test-an-EV/?EV-code=EV8>. Accessed: 2020-02.
- [TSVH21] István Táció, Bálint Sinkovics, István Vokony, and Bálint Hartmann. The challenges of low voltage distribution system state estimation—an application oriented review. *Energies*, 14(17):5363, 2021.
- [WBH<sup>+</sup>16] Ziyu Wang, Victor Bapst, Nicolas Heess, Volodymyr Mnih, Remi Munos, Koray Kavukcuoglu, and Nando de Freitas. Sample efficient actor-critic with experience replay. *arXiv preprint arXiv:1611.01224*, 2016.
- [ZMH<sup>+</sup>17] Hongcai Zhang, Scott Moura, Zechun Hu, Wei Qi, and Yonghua Song. Joint pev charging station and distributed pv generation planning. In *2017 IEEE power & energy society general meeting*, pages 1–5. IEEE, 2017.
- [ZSL<sup>+</sup>20] Hongcai Zhang, Colin JR Sheppard, Timothy E Lipman, Teng Zeng, and Scott J Moura. Charging infrastructure demands of shared-use autonomous electric vehicles in urban areas. *Transportation Research Part D: Transport and Environment*, 78:102210, 2020.

## Funding



This research was developed in the context of the “Optimization of Regional Infrastructures For The Transition To Electric And Connected Autonomous Vehicles” (EVA) project (91188), supported by the Smart Energy Systems ERA-NET program and European Union’s Horizon 2020 research and innovation programme under grant agreement no. 775970 (RegSys).