Dynamics and Optimization in Spatially Distributed Electrical Vehicle Charging

Fernando Paganini, *Fellow, IEEE,* and Andres Ferragut

***Abstract*—We consider a spatially distributed demand for electrical vehicle recharging, that must be covered by a fixed set of charging stations. Arriving EVs receive feedback on transport times to each station, and waiting times at congested ones, based on which they make a selfish selection. This selection determines total arrival rates in station queues, which are represented by a fluid state; departure rates are modeled under the assumption that clients have a given sojourn time in the system. The resulting differential equation system is analyzed with tools of optimization. We characterize the equilibrium as the solution to a specific convex program, which has connections to optimal transport problems, and also with road traffic theory. In particular a price of anarchy appears with respect to a social planner's allocation. From a dynamical perspective, global convergence to equilibrium is established, with tools of Lagrange duality and Lyapunov theory. An extension of the model that makes customer demand elastic to observed delays is also presented, and analyzed with extensions of the optimization machinery. Simulations to illustrate the global behavior are presented, which also help validate the model beyond the fluid approximation.**

***Index Terms*—Electrical vehicle charging, optimization, transportation networks, distributed algorithms/control.**

I. Introduction

The application of optimization tools to operate spatially distributed facilities has a very rich and extensive history. Perhaps the oldest such research is *optimal transport*, that goes back to Monge in the 18th century (see, e.g. [20]): the problem concerns a planner's decision on how to efficiently transport mass between demand and supply spatial distributions.

More recently, the operation of engineered *networks* often resorts to optimization, e.g. for power dispatch in the electric grid [12], or routing in communication networks [1]. These decisions must *dynamically* adapt to changing conditions and, as networks grow to enormous scale, *distributed* decisions become imperative. Designs which meet these challenges may sometimes be found through a combination of differential equations and convex optimization machinery, as was the case for resource allocation in the Internet [7], [11].

A key question in distributed network operation is: to what degree may control be imposed on individual units or, on the contrary, are these agents making their own selfish decisions? A prominent instance of the latter case is routing in road traffic networks [21]: if properly informed, drivers naturally select paths of the least latency, which results in a congestion *game*, whose equilibrium was classically analyzed by Wardrop [24]. This solution, while not centrally planned, can nevertheless be characterized in terms of a suitable optimization problem, which has been helpful to understand the *Price of Anarchy*, i.e. the gap between this equilibrium and the social welfare optimum [18], and to propose means (e.g. tolls) to mitigate it. While much of this classical analysis concerns equilibrium, *dynamic* studies of road traffic networks are also extensive, see e.g. [5] and references therein.

In this paper we consider a new application area, the operation of an Electrical Vehicle (EV) charging infrastructure. In particular, we are interested in public facilities situated in parking lots, where EV chargers are made available for temporary use [13]. This development has motivated an active area of research, within which we distinguish different problems: (i) the operation of a *single* facility of this kind, in particular the scheduling of charging opportunities taking into account EV deadlines and installation limitations [10], [25]; (ii) integration of EV charging to the smart grid [14], [23]; (iii) facility location problems, i.e. where to deploy EV charging [9], [15].

Our focus here is on the *operation* of a spatially *distributed* infrastructure made up of several charging stations, to efficiently serve a distributed demand for EV recharging. At a high level, this appears to be an optimal transport problem: demand follows a certain spatial distribution, and supply is offered in another; the optimal allocation assigns demand to stations while minimizing the overall required travel.

Charging demand does not, however, materialize in a batch; rather, we have a *dynamic* situation in which requests for service arise asynchronously over time at different spatial locations. Selecting an adequate station for each request must consider both the transport cost and station congestion. Without the transport component, such *load balancing* decisions have been studied extensively in computer networks, where routing is in charge of a central dispatcher (see e.g. [22]). Here we must incorporate the transport aspect and, importantly, compulsory routing is not assumed.

Rather, drivers will select a station to obtain the fastest possible service, similarly to selfish routing for road traffic. Indeed, our results have some parallels with this literature, but also distinguishing features. The road traffic problem considers a network of links in which all traffic is affected by selfish routing, and delays are a static function of link flows[1]. Here we are analyzing an *overlay* on the road system, in which a small portion of vehicles takes part; from this perspective,

The authors are with the Mathematics Applied to Telecommunications and Energy Research Group, Universidad ORT Uruguay, Montevideo, Uruguay (e-mail: paganini@ort.edu.uy).

Research supported by AFOSR-US under Grant FA9550-12-1-0398.

[1]Some models also include vehicle *densities* in links as variables [4], [5], and sometimes PDE effects [3]; these are beyond our scope.

transport delays are exogenous, they depend on distance and background traffic but not on EV routing decisions. On the other hand, routing will affect congestion at charging stations; we assume congestion delay information is fed back, together with transport delays, to the selfish routing agents.

Our contributions are as follows:

- We present a differential equation model for the evolution of station queues, driven by spatially distributed demand for charge and selfish decisions on station assignment. A non-standard feature is the treatment of departures: rather than require that jobs remain until completion (i.e., complete charge), for public charging facilities it is most natural to assume that EVs depart after a given *sojourn time*. This assumption impacts both queue evolution models and the calculation of queueing delays. These waiting times, superimposed to the transport delays, determine the selfish routing flows and thus the full dynamics.
- We characterize the dynamics by introducing a suitable *convex optimization*, variant of the optimal transport problem: equilibrium points are proved to correspond to optima, and global convergence to equilibrium is established. Our results differ from other such characterizations in the selfish routing literature [18], [21], due to our delay model as a function of the queue states. Proofs require extensive use of Lagrange duality, together with specialized refinements of Lyapunov-LaSalle theory.
- We compare the resulting equilibrium with the socially optimal allocation, exhibiting the gaps between the two that lead to a price of anarchy. Again, while this coincides conceptually with classical selfish routing, there are differences in the specifics.
- We extend the model to allow for *elasticity* in the demand: as a function of the experienced delays, some drivers may choose not to participate in the recharge system. These modified dynamics are also related to optimization: introducing a utility function to model customer patience, we show that the equilibrium maximizes a certain surplus objective in the input rates. Proofs require the invocation of minimax theory for a suitably chosen convex-concave function. Global convergence is also established.
- Simulation studies are carried out with a concrete instance of stations in the plane, to illustrate the resource allocation and its comparison with optimal transport. These experiments are based on discrete, stochastic demands, demonstrating the approximate validity of our model beyond the fluid abstraction.

A preliminary version of some of our results appeared in the conference paper [16]. There we used a discontinuous, switching model for selfish routing, which made the dynamic study challenging. Here we employ a smooth approximation to switching, which allows for a complete mathematical treatment within the realm of ordinary differential equations.

The rest of the paper is organized as follows. We first collect in Section I-A some notation and background material. In Section II we motivate the problem and develop our differential equation model. In Section III we present the connection to convex optimization, and the resulting interpretations. Section IV covers the version with demand elasticity. Simulations are presented in Section V, and conclusions given in Section VI. Some technical proofs are collected in the two Appendices.

A. Preliminaries and notation

We cover here some notation and background material from convex analysis. $\mathbb{R}^n$ is the standard n-dimensional space, and $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n : x_j \geq 0 \ \forall j\}$ the non-negative orthant; $\Delta_n = \{\delta \in \mathbb{R}^n_+ : \sum_{j=1}^n \delta_j = 1\}$ is the unit simplex.

Matrix variables $X = (x_{ij}) \in \mathbb{R}^{m\times n}$ are represented in uppercase, and we will use the notation $x^i = (x_{ij})_{j=1}^n \in \mathbb{R}^n$ to represent the i-th row vector of matrix X.

Both convex and concave functions $f : \mathbb{R}^n \to \mathbb{R}$ will appear; if they are differentiable, ∇f denotes the gradient. A basic non-smooth concave function $\varphi : \mathbb{R}^n \to \mathbb{R}$ is

$$\varphi(y) = \min_j(y_j) = \min_{\delta\in\Delta_n} \sum_{j=1}^n y_j\delta_j. \tag{1}$$

We will extensively use a smooth approximation to the minimum, a "log-sum-exp" function with parameter $\epsilon > 0$:

$$\varphi_\epsilon(y) := -\epsilon \log\left(\sum_j e^{-y_j/\epsilon}\right). \tag{2}$$

This function may be called a "soft-min", given the bounds:

$$\min(y_j) - \epsilon\log(n) \leq \varphi_\epsilon(y) \leq \min(y_j). \tag{3}$$

$\varphi_\epsilon(y)$ is concave, and its gradient $\nabla\varphi_\epsilon(y) =: \delta(y)$ is an element of the unit simplex Δ_n, with components

$$\delta_j(y) = \frac{e^{-\frac{y_j}{\epsilon}}}{\sum_{k=1}^n e^{-\frac{y_k}{\epsilon}}}; \tag{4}$$

these are largest for the minimizing coordinates of y.

Introduce finally the *negative entropy* function

$$\mathcal{H}(\delta) = \sum_{j=1}^n \delta_j \log(\delta_j), \quad \delta \in \Delta_n. \tag{5}$$

$\mathcal{H}$ is strictly convex, non-positive and lower bounded over the unit simplex. It is connected to log-sum-exp by convex (Fenchel) duality, as stated in the following (see e.g. [2]):

Lemma 1: For $y \in \mathbb{R}^n_+$,

$$\varphi_\epsilon(y) = \min_{\delta\in\Delta_n} \Big[\sum_{j=1}^n y_j\delta_j + \epsilon\mathcal{H}(\delta)\Big].$$

Furthermore, the unique minimizing $\delta(y)$ is given by (4).

Note that as $\epsilon \to 0+$, the optimization above approaches the one in (1), consistently with the approximation (3).

II. Dynamic model

We consider a set of EV charging *stations*, indexed by $j = 1, \ldots, n$, occupying certain locations in the territory. We will not make explicit reference to this geometry, it will be encoded in the transport costs to be specified. Each station has capacity (number of charging spots, assumed all identical) c_j.

Vehicles demanding service are distributed in space. While conceivably there could be a continuum of locations from

3

which a service request could arise, a natural simplification is to consider a discrete set of locations $i = 1, \dots m$; for instance, this could be the set of city corners, and we associate the demand location to the closest corner. Typically, $m \gg n$.

The relative positions of demand locations i and stations j are reflected in a matrix $\mathcal{K} = (\kappa_{ij})$ of transport costs. In this paper we assume cost has units of travel *time*, and is exogenous, determined by distance and background traffic. This is an appropriate assumption if the vehicles participating in recharging are a minor portion of the overall traffic.

The decision to be implemented is an *assignment* of demand to supply, i.e. of charging requests to stations, that takes into account travel costs. As a starting point for our discussion, it is useful to consider first an *optimal transport* problem[2]:

$$\min \sum_{ij} \kappa_{ij}\xi_{ij}, \text{ subject to} \tag{6}$$
$$\xi_{ij} \geq 0 \ \ \forall i,j; \quad \sum_j \xi_{ij} = \eta_i \ \ \forall i; \quad \sum_i \xi_{ij} = \sigma_j \ \ \forall j.$$

In this discrete version of the Monge-Kantorovich [20] problem[3], η_i and σ_j would be given, respectively demand and supply quantities at each location, satisfying global balance ($\sum_j \sigma_j = \sum_i \eta_i$), and one seeks a *transport plan* $\{\xi_{ij}\}$ of minimum cost. This is a *static*, one shot allocation decision.

In our application, instead, demand is *dynamic*: charging requests arrive over time, and are assigned upon arrival to a station, chosen with real-time information on the system. In addition to travel cost, current station congestion must be considered. Moreover, instead of a global planner we will have decentralized decisions consistent with selfish incentives.

In our fluid model, the rate $r_i \geq 0$ of requests/sec arising at each location i will be an exogenous quantity. Initially it will be fixed, later in Section IV we will introduce elasticity in the demand process. The model will determine a matrix variable $X(t) = (x_{ij}(t))$, representing the rates of requests from location i directed to station j, as a function of time. These will satisfy the balance conditions

$$x_{ij} \geq 0, \quad \sum_j x_{ij} = r_i \ \text{ for each } \ i = 1, \dots m,$$

of analogous form to the demand-side constraints in (6), but here in units of *rates*. It is convenient to introduce also the variables $\delta_{ij} = x_{ij}/r_i$, denoting the *fraction* of requests from location i sent to station j. For each fixed i, $\delta^i = (\delta_{ij})_{j=1}^n$ is a vector in the n-dimensional unit simplex Δ_n.

In contrast with optimal transport, here we do not fix the rates on the supply side, i.e. the EV/sec served at each station; this will be a consequence of routing decisions. If transportation costs were the only consideration, the natural assignment would be for EVs to choose the cheapest (closest in travel time) station. For a concrete visualization: if travel costs κ_{ij} are proportional to Euclidean distance, this would mean breaking up demand spatially in accordance to the Voronoi cells (see e.g. [2]) associated with station locations; Section V has illustrative pictures.

[2]Here and henceforth, for brevity we will often omit the index ranges $i = 1, \dots m$, and $j = 1, \dots n$.

[3]Also called the Hitchcock problem in the transportation literature [21].

However, station congestion should also influence the allocation decision. At each station j, denote by q_j the corresponding queue of EVs, whether or not they have a charging spot. In fluid terms, the queue dynamics are described by

$$\dot{q}_j = \sum_i x_{ij} - d_j, \quad j = 1, \dots, n; \tag{7}$$

d_j denotes the departure rate from the queue. Different models for departures could apply. In [16] we considered two options: either EVs depart when service is completed, or they depart after a certain *sojourn time*. In this paper we focus on the second option, which is most appropriate for public charging facilities, used by customers in combination with their other activities. We assume customers have a time budget for the recharge operation, and upon its expiration they will depart from the system, irrespective of the amount of service obtained; a partial recharge is still valuable. This differs from the case of home charging or charging of fleets, where a charge target would typically be required.

In practice, drivers would have different sojourn times, which could be modeled as random. For our fluid formulation, we will denote by T the mean sojourn time across the population, and use the following expression for the departure rate as a function of queue occupation:

$$d_j(q_j) = \frac{q_j}{T}. \tag{8}$$

The above is a version of Little's law in queueing theory [1], a basic conservation law between mean queues, rates and sojourn times. It is valid under very general conditions assuming balance between arrival and departure rates. Here we are applying it to departure only, allowing for a mismatch with arrivals during the transient regime. In Section V we will test this transient model through simulations.

In reference to (7), note that since $d_j(0) = 0$, and $x_{ij} \geq 0$, non-negativity of queues is automatically preserved.

Remark 1: Eq. (7) applies the input rate directly to the queues, neglecting transport delays; including them would yield a delay-differential equation. From the point of view of the *equilibrium* analysis to be carried out below, this change would be inconsequential. However it would significantly complicate the convergence analysis, hence we do not consider this effect. Our simpler model is approximately valid provided transport delays are much smaller than the time-scale of sojourn times.

If the queue q_j is below capacity, then all assigned vehicles have a corresponding slot, so there are no delays to receiving service, other than travel. On the contrary, if $q_j > c_j$ there will be an additional *waiting time*, which we proceed to model, assuming a first-come first-serve queueing policy. An arriving EV must wait for the time until the excess assignment $q_j - c_j$ is cleared by the departure process. Our model for waiting delay is thus[4]:

$$\mu_j := \frac{[q_j - c_j]^+}{d_j(q_j)} = T\left[1 - \frac{c_j}{q_j}\right]^+, \tag{9}$$

where we have invoked (8). The function $\mu_j(q_j)$ is depicted in Fig 1 below.

[4]Henceforth, $[\cdot]^+ = \max\{\cdot, 0\}$.

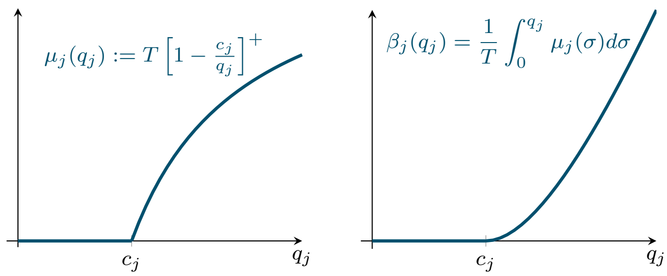

Fig. 1. Delay model (9) and penalty barrier function (12).

Consequently, an EV arriving at location i, if choosing station j, will be subject to a total delay to service of $\kappa_{ij} + \mu_j$. Note that both terms have compatible units of *time*. The main assumption that completes our model is that the information $\kappa_{ij} + \mu_j$ is available to drivers, who then make a selfish decision. Note, from a practical perspective, that time travel information is currently accessible through smartphone technology. Expanding on this, a service could be deployed through a mobile application, by which subscribers communicate bidirectionally with the stations to receive the waiting time information, and allow for stations to keep track of their customer queues.

Let $y^i = \kappa^i + \mu$ be the vector of delays observed from location i to all stations, i.e. $y^i_j = \kappa_{ij} + \mu_j,\ j = 1,\dots,n$. The natural selfish choice is $j \in \arg\min(y^i_j)$, i.e. minimizing delay to service. We analyzed this model in [16], which involves *switching*, i.e., differential equations with a discontinuous field. While some of the analysis can be carried out in this setting, convergence proofs are technically challenging.

In this paper we develop a smoother alternative, corresponding to the "soft-min" approximation described in Section I-A: routing fractions from location i will follow the expression (4), corresponding to $y^i = \kappa^i + \mu$; smaller delays are favored, but in a less drastic fashion. Such "logit" choice based on delay is a common model in the selfish routing literature (see [21]); it can be justified when there is noise in the delay information. Specifically, the routing fractions from location i are:

$$\delta_{ij}(\mathcal{K},\mu) = \frac{e^{-\frac{\kappa_{ij}+\mu_j}{\epsilon}}}{\sum_k e^{-\frac{\kappa_{ik}+\mu_k}{\epsilon}}}. \tag{10}$$

The preceding equation closes the feedback loop, leading to the overall dynamics:

$$\dot{q}_j = \sum_{i=1}^{m} x_{ij} - \frac{q_j}{T}. \quad j = 1,\dots n. \tag{11a}$$

$$\mu_j(q_j) = T\left[1 - \frac{c_j}{q_j}\right]^+, \quad j = 1,\dots n. \tag{11b}$$

$$x_{ij} = r_i\delta_{ij},\ i = 1,\dots m, j = 1,\dots n, \text{ with } \delta_{ij}(\mathcal{K},\mu) \text{ in (10).} \tag{11c}$$

The model has input parameters T, c_j, r_i, κ_{ij}. For the analysis to follow we assume constant transport costs, and thus omit henceforth the dependence on $\mathcal{K}$ in (10).

We observe first that the above differential equation has a globally Lipschitz field. Indeed, the mapping $x_{ij}(\mu)$ is continuously differentiable in $\mathbb{R}^n$, and thus admits a global Lipschitz constant over the bounded set $\mu \in [0,T)^n$, which is the range of the function $\mu(q)$ in (11b). The function $\mu_j(q_j)$ is not everywhere differentiable, but admits a global Lipschitz constant T/c_j. Thus $X(\mu(q))$ is globally Lipschitz; substitution into (11a) gives a globally Lipschitz field in q.

As a consequence, given an initial condition $q(0)$, solutions to (11) exist, are unique, and defined for all time. In the following section, we will analyze their behavior invoking tools of convex optimization.

III. Optimization characterization

We begin by introducing a barrier function, which expresses a soft version of the capacity constraints. Let:

$$\beta_j(q_j) := \int_0^{q_j} \left[1 - \frac{c_j}{\sigma}\right]^+ d\sigma \tag{12}$$

$$= \begin{cases} 0 & q_j \le c_j, \\ q_j - c_j - c_j\log\left(\frac{q_j}{c_j}\right), & q_j > c_j. \end{cases}$$

This is a convex, monotonically increasing function, also shown in Fig. 1.

We are now ready to introduce our convex optimization problem in the variables $X = (x_{ij})$ and $q = (q_j)$:

$$\min \sum_{i,j}\kappa_{ij}x_{ij} + \sum_j \beta_j(q_j) + \epsilon\sum_{i,j} x_{ij}\log\left(\frac{x_{ij}}{r_i}\right) \tag{13a}$$

$$\text{subject to:} \quad x_{ij} \ge 0\ \forall i,j; \quad \sum_j x_{ij} = r_i\ \forall i; \tag{13b}$$

$$\sum_i x_{ij} = \frac{q_j}{T}, \quad \forall j. \tag{13c}$$

Let $C(X,q)$ be the cost function in (13a). Its first term is analogous to the transportation cost in (6); without the other terms, this amounts to the optimal transport of demands r_i to supplies q_j/T, in units of (arrival and departure) rates. In contrast with (6), however, the departure rates involve free decision variables q_j, penalized by the barrier cost (12).

The other difference in (13a) is the final perturbation cost (for small $\epsilon > 0$), which can be expressed in terms of the negative entropy of the routing fractions:

$$\epsilon\sum_{i,j} r_i\delta_{ij}\log(\delta_{ij}) = \epsilon\sum_i r_i\mathcal{H}(\delta^i). \tag{14}$$

For each i, $r_i\mathcal{H}(\delta^i)$ is minimized for uniform routing fractions. Thus, including this penalty softens the selfish routing choice. In the duality-based analysis that follows, using the tools of Section I-A we show this perturbation term is directly associated with the soft-min routing choice in (10).

Remark 2: The final term in (13a) is written for $r_i > 0$; we could exclude a priori locations i with $r_i = 0$. In expression (14) these terms drop out automatically: the penalty term goes to zero as demand disappears, a feature which will be relevant for the elastic case in the following section.

Proposition 2: The optimization problem (13) is feasible, and has a unique optimal point (X^*, q^*).

Proof: For feasibility, (13b) can be satisfied by any split x_{ij} of the demands $r_i > 0$, and the free variable

5

q_j can accommodate (13c). For existence/uniqueness of the minimum, one can consider an equivalent problem in X by replacing q_j from (13c) into the cost term $\beta_j(\cdot)$. The resulting function of X is *strictly* convex (due to the negative entropy term), over a compact domain (13b): thus there is a unique minimum X^*, and consequently q^* due to (13c). ■

The main result of this section is that the optimization (13) characterizes the load balancing dynamics (11): the optimal point is a globally attractive equilibrium of the dynamics. To prove this will require the use of Lagrange duality, as follows.

A. Lagrangian and Equilibrium Characterization

Introduce the Lagrangian of Problem (13) with respect to constraints (13c):

$$\begin{aligned} L(X,q,\mu) &= \sum_{i,j}\kappa_{ij}x_{ij} + \sum_j \beta_j(q_j) + \epsilon\sum_{i,j} x_{ij}\log\left(\frac{x_{ij}}{r_i}\right) \\ &\quad + \sum_j \mu_j\left[\sum_i x_{ij} - \frac{q_j}{T}\right] \quad (15a)\\ &= \underbrace{\sum_{i,j} x_{ij}\Big[\kappa_{ij}+\mu_j+\epsilon\log\left(\frac{x_{ij}}{r_i}\right)\Big]}_{L_1(X,\mu)} \quad (15b)\\ &\quad + \underbrace{\sum_j\left[\beta_j(q_j) - \mu_j\frac{q_j}{T}\right]}_{L_2(q,\mu)}. \quad (15c)\end{aligned}$$

Suggestively, we have denoted the multipliers by μ_j; the optimum of the convex program (13) will correspond to a saddle point of this Lagrangian (minimum in (X,q), maximum in μ). We state the following result.

Theorem 3: The following are equivalent:

(i) (X^*, q^*, μ^*) is the saddle point of the Lagrangian L in (15).

(ii) (X^*, q^*, μ^*) is an equilibrium point of (11), under constant r_i.

In particular, the dynamics have a unique equilibrium point.

Proof: First observe, focusing on (15a), that for the maximum over an unconstrained μ to be finite requires primal feasibility of (13c), i.e., equilibrium of (11a):

$$\mu^* \in \arg\max_{\mu} L(X^*,q^*,\mu) \iff \sum_i x^*_{ij} = \frac{q^*_j}{T},\ j=1,\dots n. \quad (16)$$

We now look at the minimization over (X,q), treating both terms in (15b) and (15c) separately.

For the minimization of $L_1(X,\mu)$ we must consider the remaining constraints (13b) on X, which decouple across i. Invoking the routing fractions $\delta_{ij} = x_{ij}/r_i$ we write

$$L_1(X,\mu) = \sum_i r_i\left[\sum_j \delta_{ij}(\kappa_{ij}+\mu_j) + \epsilon\mathcal{H}(\delta^i)\right]. \quad (17)$$

To minimize each term in square brackets over the unit simplex, we apply Lemma 1 for the vector $y^i = \kappa^i + \mu$. The result is the log-sum-exp expression[5]

$$\varphi^i_\epsilon(\mu) := \varphi_\epsilon(\kappa^i+\mu) = -\epsilon\log\left(\sum_j e^{-(\kappa_{ij}+\mu_j)/\epsilon}\right), \quad (18)$$

achieved at $\delta^i(\mu) = (\delta_{ij}(\mu))_{j=1}^n \in \Delta_n$, which follows precisely the expression (10). We thus conclude that

$$X^* \in \arg\max_X L_1(X,\mu^*) \iff x^*_{ij} = r_i\delta_{ij}(\mu^*), \quad (19)$$
$$\text{with } \delta_{ij}(\mu) \text{ in (10).}$$

Finally, we consider the minimization of $L_2(\mu, q)$ which is unconstrained in q, and decoupled over j; we minimize each term separately. Let

$$D_{2j}(\mu_j) = \inf_{q_j}[\beta_j(q_j) - \mu_j q_j/T]. \quad (20)$$

We have the following cases:

- If $\mu_j < 0$, or $\mu_j \geq T$, $D_{2j}(\mu_j) = -\infty$.
- If $\mu_j = 0$, $D_{2j}(\mu_j) = 0$, achieved at $q_j \in (-\infty, c_j]$.
- If $0 < \mu_j < T$, the unique minimizing q_j is obtained by:

$$\beta'_j(q_j) = \left[1 - \frac{c_j}{q_j}\right] = \frac{\mu_j}{T} \iff q_j = \frac{Tc_j}{T-\mu_j}. \quad (21)$$

We can encompass the two cases with finite minimum by the relationship $\mu_j = T\left[1-\frac{c_j}{q_j}\right]^+$ between $\mu_j \in [0,T)$ and the minimizing q_j. From here we conclude that

$$q^* \in \arg\max_q L_2(q,\mu^*) \iff \mu^*_j = T\left[1-\frac{c_j}{q^*_j}\right]^+. \quad (22)$$

The left-hand sides of equations (16), (19), and (22) are the saddle point conditions (i). The corresponding right-hand sides are the equilibrium conditions (ii) for the dynamics (11).

For the last statement, since the primal variables (X^*, q^*) at a saddle point are optima, we can invoke Proposition 2 to show they are unique. Note that $\mu(q^*)$ from (11b) must be unique as well. ■

B. Dual function

As additional conclusion of the preceding analysis, let us make explicit the dual function

$$D(\mu) = \inf_{X,q} L(X,q,\mu) = \underbrace{\inf_X L_1(X,\mu)}_{D_1(\mu)} + \underbrace{\inf_q L_2(q,\mu)}_{D_2(\mu)}.$$

Referring back to (17) and (18), we have

$$D_1(\mu) = \min_{X\in(13\text{b})} L_1(X,\mu) = \sum_i r_i\varphi^i_\epsilon(\mu);$$

we further note that $D_1(\mu)$ is a differentiable function of $\mu \in \mathbb{R}^n$, and its gradient takes the form:

$$\nabla D_1(\mu) = \sum_i r_i\nabla\varphi^i_\epsilon(\mu) = \sum_i r_i\delta^i(\mu) = \sum_i x^i(\mu), \quad (23)$$

[5]This new notation emphasizes the dependence on the variable μ; the super-index i in $\varphi^i_\epsilon(\mu)$ indicates the displacement of μ by the fixed vector of transport costs κ^i from location i.

where the last expression is based on (11c).

Secondly, in reference to (21), substitution of the minimizing q_j into $[\beta_j(q_j) - \mu_j q_j/T]$ gives a minimum of

$$D_{2j}(\mu_j) = c_j \log(1 - \mu_j/T), \tag{24}$$

and this formula also covers the case $\mu_j = 0$. Therefore:

$$D_2(\mu) = \sum_j c_j \log\left(1 - \frac{\mu_j}{T}\right), \quad \text{for } 0 \le \mu_j < T. \tag{25}$$

This function is differentiable in the interior of the domain, but the boundary $\mu_j = 0$ requires some special care, as we will see in the convergence analysis below.

The overall dual function, with domain $\mu \in [0, T)^n$, is:

$$D(\mu) = \sum_i r_i \varphi_\epsilon^i(\mu) + \sum_j c_j \log\left(1 - \frac{\mu_j}{T}\right). \tag{26}$$

Proposition 4: $D(\mu)$ is strictly concave and has a finite maximum D^* over $\mu \in [0, T)^n$, achieved at a unique $\mu^* \in [0, T)^n$.

Proof: $D_1(\mu)$ is concave (not strictly), and $D_2(\mu)$ is strictly concave in $[0, T)^n$ (this follows directly from each component D_{2j}). Thus, $D(\mu)$ is strictly concave. Since $D_{2j}(\mu_j) \to -\infty$ as $\mu_j \uparrow T$, there is a global maximum $D^* = D(\mu^*)$ with μ^* strictly within $[0, T)^n$, unique due to strict concavity. This is also consistent with the uniqueness of μ^* in the saddle point, shown in Theorem 3. ■

C. Interpretation and Price of Anarchy

Equilibrium points for our model of dynamic station assignment have been shown to be solutions of a certain modified optimal transport problem. We now provide some interpretations of the result, and connections to the selfish routing literature [18], [21]. For simplicity, we will ignore in the discussion the entropy regularization term in the cost and focus on

$$C_0(X, q) := \sum_{i,j} \kappa_{ij} x_{ij} + \sum_j \beta_j(q_j); \tag{27}$$

our equilibrium (X^*, q^*) optimizes (approximately as $\epsilon \to 0$) this convex function, subject to constraints (13b)-(13c).

The first term above has a natural interpretation. Recall that κ_{ij} represents the travel times, and x_{ij} the rates, between arrival location i and station j. Therefore, $\sum_{i,j} \kappa_{ij} x_{ij}$ will be the total number of EVs currently in travel towards a charging station, a natural transport cost to be minimized.

The second term, based on $\beta_j(\cdot)$ in (12), does not have such a transparent form. From a *social welfare* perspective, the *congestion cost* to add would be the total number of EVs waiting at stations without a charging spot: $\sum_j [q_j - c_j]^+$. The natural social welfare cost would thus be:

$$C_s(X, q) = \sum_{i,j} \kappa_{ij} x_{ij} + \sum_j [q_j - c_j]^+. \tag{28}$$

We remark the following:

- $C_s(X, q)$ is also convex; in fact it is piecewise linear[6]. Using slack variables z_j, the social planner's solution to the station assignment problem can be solved by means of the linear program:

$$\min \kappa_{ij} x_{ij} + \sum_j z_j,$$

$$\text{subjet to (13b), (13c), } z_j \ge 0, \ z_j \ge q_j - c_j.$$

- $C_s(X, q) \ge C_0(X, q)$. Indeed, by (12) we have

$$C_s(X, q) - C_0(X, q) = \sum_j [q_j - c_j]^+ - \beta_j(q_j) = \sum_j c_j \log\left(\frac{q_j}{c_j}\right) \mathbf{1}_{q_j > c_j}.$$

Note that the right-hand side above is zero in the absence of congestion ($q_j \le c_j \ \forall j$). A consequence is that if the equilibrium (X^*, q^*) of our dynamics (which minimizes C_0, again ignoring the entropy term) involves no congestion, then it must also minimize $C_s(X, q)$: selfish routing achieves global welfare in this case.

However, if station congestion appears, there is a difference between both costs, the equilibrium (X^*, q^*) will no longer be socially optimal. This *Price of Anarchy* appears for similar reasons as in road traffic models [18]. To see this, rewrite the social congestion cost at station j (using (9)) as:

$$[q_j - c_j]^+ = \sum_j \left[1 - \frac{c_j}{q_j}\right]^+ q_j = \frac{1}{T} \sum_j \mu_j q_j,$$

and compare it to the corresponding barrier term in (27):

$$\beta_j(q_j) = \frac{1}{T} \int_0^{q_j} \mu_j(\sigma) d\sigma.$$

A similar variation appears in the road traffic literature, where latencies are taken to be static functions of link *flows*. The product of latency and flow is the natural welfare cost, whereas the integral of the latency function is the cost that characterizes the Wardrop equilibrium in terms of optimization.

The difference here is that the congestion delays (sensitive to routing decisions) reside at stations, not links. The state variables in our problem are station queues rather than flows, with an integrator in between. The relevant optimization problem thus differs from the classical one of Beckmann (see e.g. [19]), set purely in the domain of (primal) flow variables. Our barrier functions work with queues, and our analysis requires Lagrange duality, both for the equilibrium characterization obtained above, and for the convergence results to follow.

Remark 3: This distinction has parallels in Internet congestion control [7], [11] based on optimization. In that case, so-called "primal" models are posed entirely in terms of flow variables, whereas "dual" models include the fluid queues. The former are analyzed in the original flow variables, the latter in the space of multipliers. To our knowledge, a dual formulation has not been used so far in the context of selfish routing.

Example 1: We illustrate the Price of Anarchy through a toy example, solved numerically. There are two stations with capacities $c_1 = 20$, $c_2 = 40$, and a single source location. Transport times are $\kappa_1 = 1, \kappa_2 = 10$, sojourn time is $T = 60$. We increment the demand r upward from zero, and compare the equilibrium with the socially optimal solution.

[6]We could obtain *strict* convexity over the domain, and thus a unique optimal point for $C_s(X, q)$, by adding an entropy term as in (13a).

7

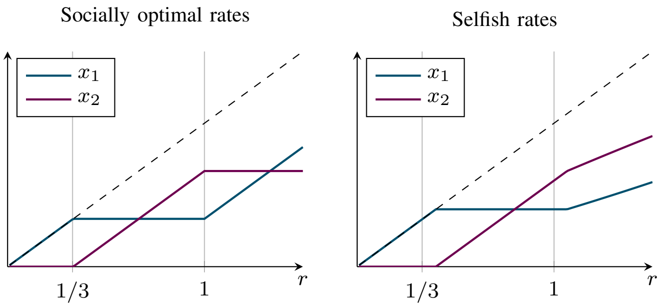

Fig. 2. Rate splitting for the socially optimal (left) and selfish (right) routing policies as a function of input rate, for Example 1.

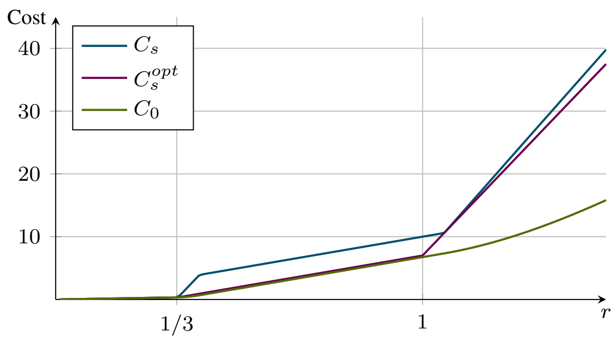

Fig. 3. Cost comparison for the socially optimal and selfish routing policies.

In this simple case, $x_i = q_i/T$ so we can write the social welfare cost as a function of rates:

$$C_s(x_1, x_2) = \kappa_1 x_1 + \kappa_2 x_2 + [Tx_1 - c_1]^+ + [Tx_2 - c_2]^+ .$$

In Fig. 2 (left) we plot the optimal breakup of rates as a function of r; initially, $x_1 = r$ so all traffic is sent to the closest station. Upon reaching $r = 1/3$ this station fills ($rT = c_1$); and after that it becomes optimal to set $x_2 = r - 1/3$, i.e. send excess traffic to station 2. This is clear from the expression above: since $\kappa_2 < T$, it is cheaper to pay the transport cost to station 2, than the congestion cost at station 1. After $r = 1$, both stations are full, so congestion cost is inevitable and also indifferent to station choice. Therefore, travel costs dictate that further increases in traffic must be sent again to station 1.

Now consider the plot (on the right) of equilibrium rates for the selfish routing case. At low loads, there is no congestion and the solution is socially optimal. After queue 1 reaches capacity, selfish routing will continue to prefer station 1 until its queueing delay μ_1 equates the difference $\kappa_2 - \kappa_1$ in transport delays. This happens at $r := r_0 \approx 0.4$; excess rates beyond this value will start choosing station 2 as in the socially optimal solution; but the waiting cost $\mu_1 \cdot r_0$ implies inefficiency. Station 2 congests at $r = r_0 + 2/3$, after which selfish routing spreads the additional increase between both stations; this is required to maintain the indifference condition $\kappa_1 + \mu_1 = \kappa_2 + \mu_2$, and again implies some inefficiency.

Fig. 3 shows, as a function of r: the cost C_0 characterizing the selfish equilibrium, the resulting social cost C_s, and the optimum social cost C_s^{opt}. There is no price of anarchy at low loads; in the intermediate region inefficiency appears, for the most part constant at $(\kappa_2 - \kappa_1)r_0$. There is a single point of efficiency again when both stations reach congestion, after which we observe a roughly linear gap in cost (constant price of anarchy) between both solutions.

D. Convergence

Beyond the characterization of the equilibrium, we will establish that it is globally attractive under the dynamics (11). From the Lipschitz nature of the field we know that from any initial condition $q(0)$, the trajectory $q(t)$ is well-defined for all time, and continuously differentiable. From this we can introduce the functions $\mu(q(t))$ and $D(\mu(q(t)))$; a key step of the convergence analysis will be to prove the latter is monotonically increasing along trajectories.

These composite functions are not, however, differentiable. In particular the formula $\mu_j(q_j) = T\,[1 - c_j/q_j]^+$ is not differentiable at $q_j = c_j$ (see Fig. 1). Similarly, if we compose it with $D_{2j}(\mu_j)$ in (24) we obtain:

$$D_{2j}(\mu_j(q_j)) = \begin{cases} 0, & q_j \le c_j; \\ c_j \log\left(\frac{c_j}{q_j}\right), & q_j > c_j; \end{cases} \qquad (29)$$

this (non-positive) function is not differentiable at $q_j = c_j$.

Nevertheless, both $\mu_j(q_j)$ and $D_{2j}(\mu_j(q_j))$ are Lipschitz functions; composing them with the smooth $q(t)$ will yield *absolutely continuous* functions $\mu_j(q_j(t))$ and $D_{2j}(\mu_j(q(t)))$; time derivatives exist almost everywhere, and the functions are integrals in time of their derivatives. We will denote by $\dot{\mu}_j$ and $\dot{D}_{2j}$ these derivatives, and let $\mathcal{T} \subset \mathbb{R}_+$ be the set of times for which they are well-defined for every j. The complement of $\mathcal{T}$ has Lebesgue measure zero.

Lemma 5: For every $t \in \mathcal{T}$, we have

$$\dot{D}_{2j}(t) = \frac{-c_j}{T - \mu_j} \dot{\mu}_j(t); \qquad (30)$$

furthermore, if $q_j(t) \le c_j$, then $\dot{\mu}_j = \dot{D}_{2j} = 0$.
Proof is given in Appendix A. We now get to the monotonicity result.

Proposition 6: Consider a trajectory $q(t)$ of (11). For any $t \in \mathcal{T}$ defined above, the dual function (26) satisfies:

$$\frac{d}{dt} D(\mu(q(t))) = \sum_{j: q_j(t) > c_j} T c_j \left[\frac{\dot{q}_j}{q_j}\right]^2 \ge 0. \qquad (31)$$

Consequently, $D(\mu(q))$ is non-decreasing along trajectories $q(t)$ arising from the dynamics (11).

Proof: We differentiate separately each term of the dual function; since $D_1(\mu)$ is differentiable we apply the chain-rule and the gradient in (23) to obtain:

$$\dot{D}_1 = \langle \nabla D_1(\mu), \dot{\mu} \rangle = \sum_i \langle x^i(\mu), \dot{\mu} \rangle = \langle \dot{q} + q/T, \dot{\mu} \rangle;$$

here $\langle \cdot, \cdot \rangle$ is the standard inner product in $\mathbb{R}^n$, and we have invoked (11a). Note that $\dot{\mu}$ is well-defined within $\mathcal{T}$.

Now consider $D_2(\mu)$, and apply Lemma 5 to obtain:

$$\dot{D}_2 = \sum_j \dot{D}_{2j} = \sum_j \frac{-c_j}{T - \mu_j} \dot{\mu}_j = \sum_{j: q_j(t) > c_j} \frac{-q_j}{T} \dot{\mu}_j = -\langle q/T, \dot{\mu} \rangle.$$

For the third step we have used two facts: first, according to Lemma 5 we may ignore in the sum terms where $q_j(t) \le c_j$, since $\dot{\mu}_j = 0$ in that case. Second, for j where $q_j(t) > c_j$, we substitute $\mu_j = T(1 - c_j/q_j)$ from (11b).

Adding both derivatives and cancelling terms we find

$$\dot{D} = \langle \dot{q}, \dot{\mu} \rangle = \sum_j \dot{q}_j \dot{\mu}_j = \sum_{j: q_j(t) > c_j} \dot{q}_j \mu_j'(q_j) \dot{q}_j$$
$$= \sum_{j: q_j(t) > c_j} \frac{T c_j}{q_j^2} [\dot{q}_j]^2; \quad t \in \mathcal{T}.$$

Again we have removed terms with $q_j \le c_j$, and in the rest we may apply the chain rule. This establishes (31) almost everywhere. Integrating in time and invoking absolute continuity, D must be non-decreasing along trajectories. ∎

Since $D(\mu(q(t)))$ is non-decreasing and bounded above, it must approach a finite limit as $t \to \infty$. We will show that this limit must be the global maximum $D(\mu^*)$, and our trajectory converges to the unique equilibrium. We state the corresponding result.

Theorem 7: Any trajectory $(q(t), X(t), \mu(t))$ of (11) converges asymptotically to the equilibrium point (X^*, q^*, μ^*).

The proof is based on a Lyapunov-LaSalle type argument [8] with the function $D(\mu(q))$, exploiting (31) to characterize points where $\dot{D} = 0$. Some non-standard modifications are required, details are given in Appendix B.

IV. Elastic demand

In this section we add a new ingredient to the formulation: instead of having a rigid quantity r_i of EVs per second seeking charge from each location i, we consider the fact that demand itself is sensitive to the delays of the installation. Namely, if τ_i is the delay that EVs from location i experience before receiving service, some customers may not be willing to wait this much and desist from entering the system.

To model this phenomenon, assume that arriving customers at location i have a willingness-to-wait characteristic, which is a random variable T_i with complementary cumulative distribution function $p_i(\tau_i) = \mathcal{P}(T_i > \tau_i)$. T_i is assumed to be absolutely continuous, so $p_i(\tau_i)$ is continuous, decreasing from 1 to 0 as τ_i moves in the range $[0, \infty)$.

Randomness reflects customer heterogeneity; now, for a deterministic fluid model, we will not incorporate stochasticity at the individual EV level but rather consider its effect in large numbers. In this sense, if $\bar{r}_i$ is the (fixed) maximum rate of requests originating at i, this rate will be "thinned" by the probability $p_i(\tau_i)$, giving

$$r_i(\tau_i) = \bar{r}_i p_i(\tau_i). \tag{32}$$

Example 2: Assume that T_i is a uniform random variable in the interval $[0, T]$, where T is the sojourn time. In that case we have the linear thinning $r_i(\tau_i) = \bar{r}_i \left[1 - \tau_i/T\right]^+$, that shuts down demand as τ_i reaches T.

Remark 4: While this method accommodates heterogeneity in customer willingness-to-wait, we are not making explicit the heterogeneity of sojourn times. In practice, these two features might have correlation as customers with more urgent charging needs must accept longer sojourn and waiting times. To account for such correlations requires a stochastic approach which is currently beyond our scope.

It will be useful to interpret (32) as a *demand curve* that specifies the quantity r_i as a function of the "price" τ_i. In microeconomic language, this is equivalent to introducing an increasing, concave *utility function* $U_i(r_i)$, and stating that the rate is chosen according to

$$r_i = \arg\max_{r_i \ge 0} [U_i(r_i) - \tau_i r_i]. \tag{33}$$

Assuming differentiability, the above amounts to the first order condition $U_i'(r_i) = \tau_i$; this should be the inverse function of (32).

Example 3: Continuing with the uniform example, we find the corresponding utility. $U_i'(r_i)$ must be linear, decreasing from T to zero in the interval $[0, \bar{r}_i]$. Therefore:

$$U_i(r_i) = \begin{cases} T r_i \left(1 - \frac{r_i}{2\bar{r}_i}\right) & 0 \le r_i \le \bar{r}_i; \\ T \frac{\bar{r}_i}{2}, & r_i > \bar{r}_i. \end{cases}$$

Note that utility will saturate (demand will satiate) at $r_i = \bar{r}_i$; to simplify matters we will assume the following property, which holds in the example above:

Assumption 1: $U_i(r_i)$ is *strictly* concave within the interval $r_i \in [0, \bar{r}_i]$.

A. Dynamics under elastic demand

We now incorporate the elastic demand feature into the framework of Section II. We need to specify the appropriate τ_i for which to apply the expression (32). Under selfish routing to the station(s) with smallest travel + waiting time, the delay-to-service experienced would be

$$\tau_i = \min_j (\kappa_{ij} + \mu_j) = \varphi(\kappa^i + \mu),$$

invoking the notation (1). Once again, for smoothness reasons we will use instead the log-sum-exp approximation (2) to the minimum, namely:

$$\tau_i(\mu) = \varphi_\epsilon(\kappa^i + \mu) = \varphi_\epsilon^i(\mu), \tag{34}$$

recalling the notation (18). We are now ready to formulate the new dynamic model, incorporating into (11) the elastic rate thinning in (32):

$$\dot{q}_j = \sum_{i=1}^m x_{ij} - \frac{q_j}{T}, \quad j = 1, \dots n. \tag{35a}$$

$$\mu_j(q_j) = T \left[1 - \frac{c_j}{q_j}\right]^+, \quad j = 1, \dots n. \tag{35b}$$

$$x_{ij} = r_i \delta_{ij}(\mu), \ i = 1, \dots m, \ j = 1, \dots n, \text{ with } \delta_{ij}(\mu) \text{ in (10)}; \tag{35c}$$

$$r_i = \bar{r}_i p_i(\varphi_\epsilon^i(\mu)). \tag{35d}$$

As in the case of (11), by substitution the above dynamics is equivalent to a differential equation $\dot{q} = f(q)$ with a globally Lipschitz right-hand side; the assumptions on $p_i(\cdot)$ preserve this feature. Hence, we again have unique solutions, defined for all time. We will also analyze the new dynamics through convex optimization.

9

B. Equilibrium characterization

We reconsider the cost function in (13a), denoted here by

$$C(X,q,r) := \sum_{i,j} \kappa_{ij} x_{ij} + \sum_j \beta_j(q_j) + \epsilon \sum_{i,j} x_{ij} \log\left(\frac{x_{ij}}{r_i}\right).$$

As compared to Section III, we now include as *variable* the vector of demand rates r, with domain $R := \Pi_{i=1}^n [0, \bar{r}_i]$; zero rates are included, in this regard we refer to Remark 2.

$C(X,q,r)$ is *not* a jointly convex function with this new variable, due to the entropy term; nevertheless, convexity holds for the solution of Problem (13) as a function of r.

Proposition 8: Let $\mathcal{C}(r)$ be the minimum of $C(X,q,r)$ over (X,q) satisfying (13b)-(13c) . Then $\mathcal{C}(r)$ is convex.

Proof: From our analysis of duality in the previous section, for each r we have $\mathcal{C}(r) = \max_\mu D(r,\mu)$, where

$$D(r,\mu) = \sum_i r_i \varphi_\epsilon^i(\mu) + \sum_j c_j \log\left(1 - \frac{\mu_j}{T}\right),$$

the dual cost in (26), making explicit its (linear) dependence on r. The maximum over μ of this family of linear functions is convex in r. ■

The optimal cost for a vector of demand rates is naturally combined with the total associated utility:

$$\psi(r) := \mathcal{C}(r) - \sum_i U_i(r_i). \tag{36}$$

Minimizing the convex function $\psi(r)$ is equivalent to maximizing $-\psi(r)$, a *social suprlus* (utility minus cost); we will show that our dynamics converges to such optimal demand rates. To pursue the analysis, it is convenient to introduce

$$\begin{aligned} W(r,\mu) :&= D(r,\mu) - \sum_i U_i(r_i), \qquad (37)\\ &= \sum_i [r_i \varphi_\epsilon^i(\mu) - U_i(r_i)] + \sum_j c_j \log\left(1 - \frac{\mu_j}{T}\right), \end{aligned}$$

defined for $(r,\mu) \in R \times [0,T)^n$. Note the following properties:

- For fixed r, $W(r,\mu)$ is strictly concave in $\mu \in [0,T)^n$. $W(r,\mu) \to -\infty$ when $\mu_j \uparrow T$. From Proposition 8 we also obtain: $\psi(r) = \max_\mu W(r,\mu)$.
- For fixed μ, $W(r,\mu)$ is convex in $r \in R$; under Assumption 1, it is strictly convex.

We now identify the saddle points of $W(r,\mu)$.

Proposition 9: There exists a unique $(r^*,\mu^*) \in R \times [0,T)^n$, such that

$$W(r^*,\mu) \le W(r^*,\mu^*) \le W(r,\mu^*) \quad \forall r \in R, \mu \in [0,T)^n.$$

Furthermore, $r^* = \arg\min \psi(r)$.

Proof: Existence of a saddle point for a convex-concave function is a minimax result, of which there are multiple versions; in this case with bounded domains it can be found in [17] (Corollary 37.6.1). Uniqueness of the saddle point follows from strict concavity/convexity. For the final statement, note:

$$\begin{aligned} \psi(r^*) &= \max_\mu W(r^*,\mu) = W(r^*,\mu^*)\\ &\le W(r,\mu^*) \le \max_\mu W(r,\mu) = \psi(r). \end{aligned}$$

■

We are now ready for our main result on equilibrium:

Theorem 10: The following are equivalent:

(i) (r^*,μ^*) is the saddle point of W in (37), and given r^*, (X^*,q^*) are the solution to the optimization (13).

(ii) (X^*,q^*,μ^*,r^*) is an equilibrium point of (35).

In particular, the dynamics have a unique equilibrium point.

Proof: Start from (i). At the saddle point (r^*,μ^*) we have $r^* = \arg\min_r W(r,\mu^*)$. From (37) we see that

$$\begin{aligned} r_i^* &= \arg\min_{r_i} [r_i \varphi_\epsilon^i(\mu^*) - U_i(r_i)]\\ &= \arg\max_{r_i} [U_i(r_i) - r_i \varphi_\epsilon^i(\mu^*)] = \bar{r}_i p_i(\varphi_\epsilon^i(\mu^*)); \end{aligned}$$

in the last step we used the characterization of the utility function in (33). This equation is consistent with (35d). Given r_i^*, we can invoke Theorem 3 to identify the saddle conditions for (q^*,X^*,μ^*) with the equilibrium of (35a)-(35c). Therefore we have (ii).

Now start with (ii). For the given r^*, we know from Theorem 3 that (q^*,X^*,μ^*) are primal-dual optimal for the optimization (13); in particular μ^* is dual optimal,

$$\mu^* = \arg\max_\mu D(r^*,\mu) = \arg\max_\mu W(r^*,\mu).$$

Finally, (35d) implies $r^* = \arg\min_r W(r,\mu^*)$ (these steps are reversible). Therefore (r^*,μ^*) is the saddle point of W, and we have (i).

Uniqueness of the equilibrium follows from the uniqueness of (r^*,μ^*) and the application of Theorem 3 for the remaining variables. ■

The previous result provides an interpretation of the equilibrium of the elastic dynamics, extending the one from Section III-C. Indeed, if we consider jointly the two statements in condition (i), what the equilibrium achieves is the minimization of the cost $C(X,q,r) - \sum_i U_i(r_i)$ over all available degrees of freedom. Thus the dynamics achieve a specific welfare optimization, in which:

- The utilities in question are directly related to the elastic thinning rule in (32).
- The cost portion $C(X,q,r)$ carries, again, the natural cost of transportation, a (negligible) regularization term, and the barrier terms $\beta_j(q_j)$. As in Section III-C, these exhibit a deviation with the respect to the waiting cost. Thus there may be inefficiency with respect to the natural social welfare optimization under elastic demand.

C. Convergence

Again, we show that our equilibrium is globally attractive under the dynamics (35). The analysis parallels Section III-D, based on the monotonicity of a certain function along trajectories of our dynamics. Here the function in question will be

$$\begin{aligned} \mathcal{D}(\mu) &:= \min_r W(r,\mu) = \mathcal{D}_1(\mu) + D_2(\mu) \qquad (38)\\ &= \sum_i U_i^*(\varphi_\epsilon^i(\mu)) + \sum_j D_{2j}(\mu_j). \end{aligned}$$

For the first term above we have introduced the Fenchel conjugate

$$U_i^*(\tau_i) := \min_{r_i} [\tau_i r_i - U_i(r_i)] \tag{39}$$

of the concave utility function [17]. The above minimum is achieved at $r_i(\tau_i) = \bar{r}_i p_i(\tau_i)$. Since the latter function is continuous, it is not hard to show that $U_i^*(\tau_i)$ is differentiable and non-decreasing in $\tau_i > 0$, with derivative $r_i(\tau_i)$. In (38), this function is composed with the smooth function $\varphi_\epsilon^i(\mu)$. Therefore the first term $\mathcal{D}_1$ in (38) is differentiable in μ, with

$$\nabla \mathcal{D}_1(\mu) = \sum_i \bar{r}_i p_i(\varphi_\epsilon^i(\mu))\delta^i(\mu) = \sum_i x^i(\mu),$$

where we have substituted expressions from the dynamics (35); this formula is completely analogous to (23) for the inelastic demand case. The second term $D_2(\mu)$ in (38) is identical to (26). Therefore we are in a position to replicate the analysis of the previous section.

In particular, $\mathcal{D}(\mu(q(t)))$ will be absolutely continuous along trajectories of the dynamics, and we can identify a set $\mathcal{T}$ of times (whose complement has zero Lebesgue measure) where Lemma 5 holds, as well as the following extension of Proposition 6. The proof is analogous and is omitted.

Proposition 11: Consider a trajectory $q(t)$ of (35). For any $t \in \mathcal{T}$ defined above, the dual function (38) satisfies:

$$\frac{d}{dt}\mathcal{D}(\mu(q(t))) = \sum_{j: q_j(t) > c_j} T c_j \left[\frac{\dot{q}_j}{q_j}\right]^2 \geq 0. \qquad (40)$$

Consequently, $\mathcal{D}(\mu(q))$ is non-decreasing along trajectories $q(t)$ arising from the dynamics (35).

Note, from the strict concavity of $W(r, \mu)$ in μ, that $\mathcal{D}(\mu)$ in (38) is strictly concave. Also, from the min-max inequality

$$\begin{aligned}\max_\mu \mathcal{D}(\mu) &= \max_\mu \min_r W(r, \mu) \\ &\leq \min_r \max_\mu W(r, \mu) = \min_r \psi(r) = \psi(r^*),\end{aligned}$$

we conclude that $\mathcal{D}(\mu)$ is upper bounded by the optimal cost. There is equality above at the (unique) saddle point (r^*, μ^*), so we conclude that $\mathcal{D}(\mu)$ achieves its maximum (only) at μ^*. These are analogous properties to the ones we had for $D(\mu)$ in the previous section, leading to following convergence result.

Theorem 12: Any trajectory $(q(t), X(t), \mu(t), r(t))$ of (35) converges asymptotically to the equilibrium point (X^*, q^*, μ^*, r^*) characterized in Theorem 10.

The proof is very similar to the one for Theorem 7, and is omitted due to space limitations.

V. Stochastic Simulations

In this section we present simulations of the selfish load balancing dynamics. A first objective is illustrating in a concrete geometric example the properties of the equilibrium (optimum of Problem (13)), and the convergence of trajectories. A separate purpose is validating our fluid model as representative of more realistic conditions: in particular, we will simulate a *stochastic* system generated by EVs arriving randomly in time and space, with random sojourn times, and stations assigned through selfish routing. We also include delays, distinguishing the event of arrival into the system from the arrival at the station queues. We focus on the inelastic (fixed demand) case.

Our spatial domain is a square region, where recharge requests arrive as a Poisson process of overall rate $r = 3$ EVs/min, spawning at a random spatial location with uniform distribution. Sojourn times are independent and exponentially distributed with mean $T = 90$ min., so the total number of EVs present in steady-state is $rT = 270$ on average. We fix 5 charging stations with $c_j = 50$ slots each, at random positions. Since $rT > \sum_j c_j$, there is overall congestion; it will affect stations asymmetrically due to their locations.

Travel times κ_{ij} are modeled as the Euclidean distance divided by a speed v, chosen such that the travel time horizontally across the region is 50 minutes. With the given station locations, the maximum time to reach the closest station is approximately 30 min. On the left side of Fig. 4 we plot the station positions as well as the Voronoï cells that break up the plane if we use the closest station criterion.

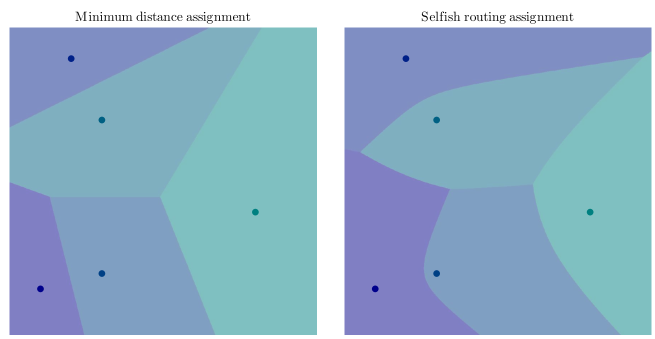

Fig. 4. Charging station positions, minimum distance cells and attraction regions in equilibrium.

In Fig. 5 we plot the station occupation resulting from our spatial stochastic simulation, using the Julia library EVQueues [6]. The selfish policy is applied, so that vehicles upon arrival are directed to the station that provides minimum delay to service. We also plot for comparison the computed numerical solutions of the fluid model (11): since the model uses a finite set of arrival locations, we discretize the space to a uniform grid of 10000 points. Note that the fluid model correctly captures both the transient and steady-state behavior of the stochastic system, and converges to the predicted equilibrium, the solution of Problem (13), shown with dotted lines.

Initially, all stations are uncongested and EVs route themselves to the closest station. Station load is thus proportional to the area of the respective Voronoï cells, so it is asymmetric. In particular the easternmost station (labeled by 1) gets congested first: as its queueing delay builds up, neighboring stations start to receive more traffic; inflection points in the respective fluid trajectories signal this event. As the simulation progresses other stations reach congestion in succession, except for Station 5 (southwest point) which stays below capacity.

In steady state, Station 1 operates with $q_1 \approx 62$, i.e. 12 requests in waiting, a queueing delay of $\mu_1 \approx 17.4$ min. The remaining congested stations reach queueing delays $\mu_2 \approx 8.9$, $\mu_3 \approx 6.4$, and $\mu_4 \approx 5.4$ minutes, while μ_5 remains at 0.

The effect of these delays in steady-state routing is observed by plotting the attraction regions, on the right of Fig. 4. Congested stations see their regions shrink, while less congested stations cover additional ground to compensate. Indifference curves can be shown to be arcs of hyperbolas in this case.

11

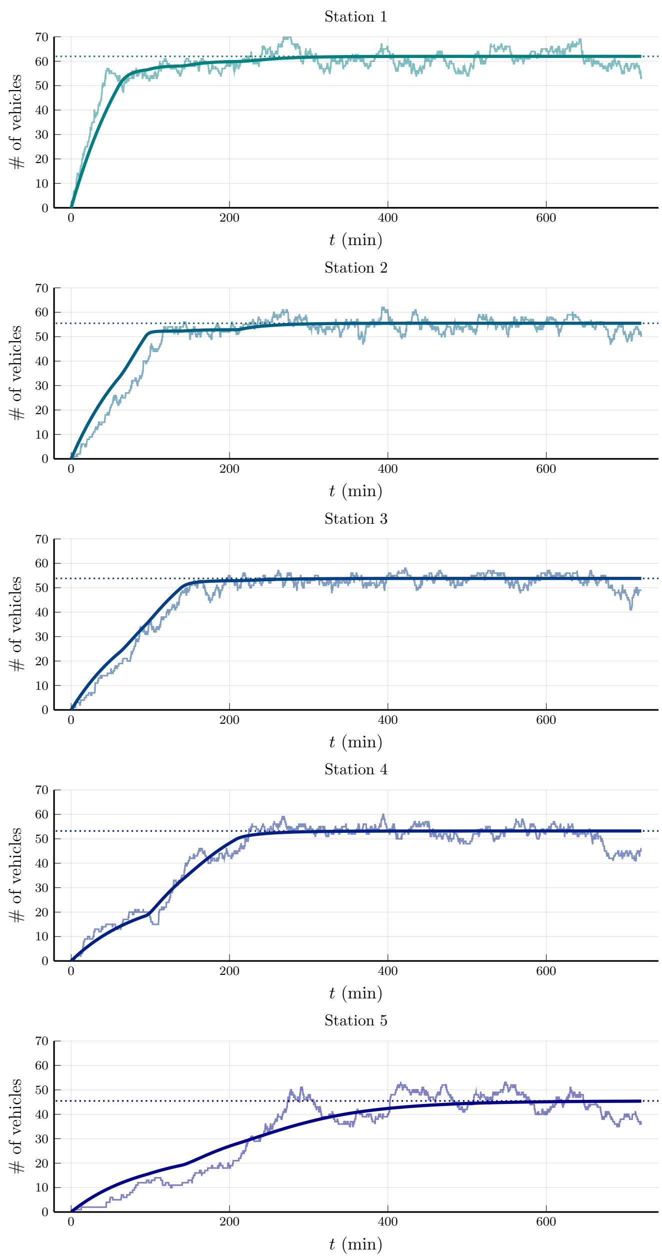

Fig. 5. Time evolution of station occupations for the stochastic system and fluid approximation (thick lines). The predicted equilibrium of Theorem 3 is shown in dotted lines.

VI. Conclusions

Network resource allocation calls for an interplay between dynamics and optimization. Besides helping design *control* mechanisms for engineered systems with centrally adjudicated resources, optimization has also been applied successfully to characterize congestion *games* between selfish agents.

In this paper we have analyzed the operation of spatially distributed EV charging resources. Rather than a static, centrally planned optimal transport we have considered dynamic, selfish assignment by EV drivers endowed with delay information. We have proposed a dynamic model and analyzed its equilibrium and dynamics with tools of convex optimization.

Natural lines for future research are: (i) Quantitative assessment of the price of anarchy, and its possible mitigation through more active control; (ii) Dynamics beyond constant demand (transients, tracking of daily variations); (iii) Stochastic versions of the queueing model and/or the geometry, for a more thorough analysis of variability.

Appendix A
Proof of Lemma 5

We must establish the identity (30) for functions $\mu_j(q_j(t))$ and $D_{2j}(\mu_j(q_j(t)))$ at times when they are both differentiable. This includes all $t : q_j(t) \neq c_j$, since at these points $\mu_j(q_j)$ from (9) and $D_{2j}(\mu_j(q_j))$ from (29) are both differentiable, and so is $q_j(t)$. In this case, (30) just follows from the chain rule, recalling that $D_{2j}(\mu_j) = c_j \log(1-\mu_j/T)$. Furthermore, if $q_j(t) < c_j$ both sides of (30) are zero.

It remains to consider the case $q_j(t) = c_j$, where $\mu_j(q_j)$ and $D_{2j}(\mu_j(q_j))$ are not smooth, they have finite but unequal lateral derivatives. Still, we are assuming $t \in \mathcal{T}$ so the composite functions $\mu_j(q_j(t))$ and $D_{2j}(\mu_j(q_j(t)))$ are differentiable. For this to happen their derivatives must be zero, as follows from a basic fact from differentiation, stated as Lemma 13 below (proof omitted for brevity). So (30) holds in this case too. ∎

Lemma 13: Let $f : \mathbb{R} \to \mathbb{R}$ be continuous, $f(c) = 0$ and with different lateral derivatives $f'(c-) \neq f'(c+)$. Let $q : \mathbb{R} \to \mathbb{R}$ be differentiable, $q(t_0) = c$. If the composite function $g(t) = f(q(t))$ is differentiable at t_0, then $\dot{q}(t_0) = \dot{g}(t_0) = 0$.

Appendix B
Proof of Theorem 7

The function $V(q) = D(\mu(q))$ is well-defined over $q \in \mathbb{R}^n$, and bounded, with a maximum $V^* = D(\mu^*)$ achieved at $q = q^*$, the equilibrium queues. Furthermore, $V(q)$ is non-decreasing along trajectories $q(t)$ of the dynamics (11). These elements suggest using V as a Lyapunov function to establish convergence. A few difficulties arise, however: when some $\mu_j^* = 0$, it is not true that $V(q) = V^*$ *only* at equilibrium q. Furthermore, we have to deal with non-differentiability. Thus, we develop a specialized variant of the LaSalle principle [8].

We begin by noting that the dynamics in $q(t)$ have a compact positively invariant set $[0, rT]^n$, where we denote $r = \sum_i r_i$. Indeed, since (11c) implies $x_{ij}(t) < r_i \ \forall \ t$, we obtain from (11a) the inequality

$$T\dot{q}_j < \sum_i Tr_i - q_j = Tr - q_j.$$

For $q_j \geq rT$ we have $\dot{q}_j < 0$; this implies a trajectory starting from $q_j \in [0, rT]$ cannot exceed the upper limit. In fact, a slight refinement of this argument implies that from any initial condition outside $[0, rT]^n$, this set is reached in finite time. Hence, it suffices to analyze the dynamics with initial conditions within this invariant set.

Consider $q(0) \in [0, rT]^n$, and the resulting trajectory $q(t)$. From Proposition 6 we know that $V(q(t))$ is monotonically non-decreasing, let its limit be $\bar{V}$.

Let L^+ be the Ω-limit set of the trajectory, itself an invariant set under the dynamics. By continuity, $V(q) \equiv \bar{V}$ for $q \in L^+$.

Now consider an auxiliary trajectory, $\tilde{q}(t)$, with initial condition $q^+ \in L^+$. We conclude that

$$V(\tilde{q}(t)) = D(\mu(\tilde{q}(t))) \equiv \bar{V} \implies \frac{d}{dt}D(\mu(\tilde{q}(t))) \equiv 0.$$

In reference to (31), this implies that $\dot{\tilde{q}}_j = 0$ almost everywhere, for any $j : \tilde{q}_j(t) > c_j$; these congested queues must be at equilibrium, and will satisfy $\dot{\mu}_j \equiv 0$. But non-congested queues ($\tilde{q}_j(t) \leq c_j$) also satisfy $\dot{\mu}_j = 0$ by Lemma 5; therefore $\mu(\tilde{q}(t)) \equiv \tilde{\mu}$, constant. Consequently, the rates $X(\tilde{\mu})$ defined by (11c) are also constant.

Stations with $\tilde{\mu}_j = 0$ ($\tilde{q}_j(t) \leq c_j$) might not be at equilibrium; however they receive a constant input rate $\sum_i x_{ij}(\tilde{\mu})$ and evolve according to the first-order linear dynamics

$$\dot{\tilde{q}}_j = \sum_i x_{ij}(\tilde{\mu}) - \tilde{q}_j/T.$$

This ODE has solutions that converge exponentially to $q_j^* = T\sum_i x_{ij}(\tilde{\mu}) \in [0, c_j]$, an equilibrium. Thus, all queues reach equilibrium in the trajectory $\tilde{q}(t)$; since the equilibrium is unique from Theorem 3, we conclude from this exercise that necessarily $\tilde{\mu} = \mu^*$, the dual optimal price, and $\bar{V} = V^* = D(\mu^*)$.

Return now to the *original* trajectory $q(t)$. We know that $V(q(t)) = D(\mu(q(t))) \to D(\mu^*)$; since $D(\mu)$ is strictly concave, with a unique maximum, we must have $\mu(q(t)) \to \mu^*$.

Consequently, the input rate $\sum_i x_{ij}(\mu(q(t))$ to the j-th station converges to the optimal rate $\sum_i x_{ij}^*$. For the stable first order system (11a), it follows that the output $q_j(t)$ converges to the resulting equilibrium: $q_j(t) \to q_j^*$ for all j. ■

References

[1] D. Bertsekas and R. Gallager, *Data networks*. Athena Scientific, 2021.
[2] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.
[3] G. Coclite, M. Garavello, B. Piccoli, *et al.*, "Traffic flow on a road network," *SIAM JOURNAL ON MATHEMATICAL ANALYSIS*, vol. 36, no. 6, pp. 1862–1886, 2005.
[4] G. Como, E. Lovisari, and K. Savla, "Throughput optimality and overload behavior of dynamical flow networks under monotone distributed routing," *IEEE Transactions on Control of Network Systems*, vol. 2, no. 1, pp. 57–67, 2015.
[5] G. Como and R. Maggistro, "Distributed dynamic pricing of multiscale transportation networks," *IEEE Transactions on Automatic Control*, vol. 67, no. 4, pp. 1625–1638, 2022.
[6] A. Ferragut, "`EVQueues.jl`: A Julia simulator of ev charging policies." `https://github.com/aferragu/EVQueues.jl`, 2021.
[7] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan, "Rate control for communication networks: shadow prices, proportional fairness and stability," *Journal of the Operational Research society*, vol. 49, pp. 237–252, 1998.
[8] H. K. Khalil, "Nonlinear systems," *Upper Saddle River*, 2002.
[9] M. Kuby and S. Lim, "The flow-refueling location problem for alternative-fuel vehicles," *Socio-Economic Planning Sciences*, vol. 39, no. 2, pp. 125–145, 2005.
[10] Z. J. Lee, G. Lee, T. Lee, C. Jin, R. Lee, Z. Low, D. Chang, C. Ortega, and S. H. Low, "Adaptive charging networks: A framework for smart electric vehicle charging," *IEEE Transactions on Smart Grid*, vol. 12, no. 5, pp. 4339–4350, 2021.
[11] S. H. Low, F. Paganini, and J. C. Doyle, "Internet congestion control," *IEEE control systems magazine*, vol. 22, no. 1, pp. 28–43, 2002.
[12] J. A. Momoh, *Electric power system applications of optimization*. CRC press, 2017.
[13] M. Mozaffari, H. A. Abyaneh, M. Jooshaki, and M. Moeini-Aghtaie, "Joint expansion planning studies of EV parking lots placement and distribution network," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 10, pp. 6455–6465, 2020.
[14] J. C. Mukherjee and A. Gupta, "A review of charge scheduling of electric vehicles in smart grid," *IEEE Systems Journal*, vol. 9, no. 4, pp. 1541–1553, 2014.
[15] F. Paganini, E. Espíndola, D. Marvid, and A. Ferragut, "Optimization of spatial infrastructure for EV charging," in *61st IEEE Conference on Decision and Control*, 2022.
[16] F. Paganini and A. Ferragut, "Dynamic load balancing of selfish drivers between spatially distributed electrical vehicle charging stations," in *59th Allerton Conference*, 2023.
[17] R. T. Rockafellar, "Convex analysis," 1970.
[18] T. Roughgarden, *Selfish routing and the price of anarchy*. MIT press, 2005.
[19] T. A. Roughgarden, *Selfish routing*. PhD Dissertation, Cornell University, 2002.
[20] F. Santambrogio, "Optimal transport for applied mathematicians," *Birkäuser, NY*, vol. 55, no. 58-63, p. 94, 2015.
[21] Y. Sheffi, *Urban transportation networks*. Prentice-Hall, Englewood Cliffs, NJ, 1985.
[22] M. Van der Boor, S. C. Borst, J. S. Van Leeuwaarden, and D. Mukherjee, "Scalable load balancing in networked systems: A survey of recent advances," *SIAM Review*, vol. 64, no. 3, pp. 554–622, 2022.
[23] Q. Wang, X. Liu, J. Du, and F. Kong, "Smart charging for electric vehicles: A survey from the algorithmic perspective," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 1500–1517, 2016.
[24] J. G. Wardrop, "Some theoretical aspects of road traffic research." *Proceedings of the Institution of Civil Engineers*, vol. 1, no. 3, pp. 325–362, 1952.
[25] M. Zeballos, A. Ferragut, and F. Paganini, "Proportional fairness for EV charging in overload," *IEEE Transactions on Smart Grid*, vol. 10, no. 6, pp. 6792–6801, 2019.

Fernando Paganini (M'90–SM'05–F'14) received his degrees in both Electrical Engineering and Mathematics from Universidad de la República, Montevideo, Uruguay, in 1990, and his M.S. and PhD degrees in Electrical Engineering from the California Institute of Technology, Pasadena, in 1992 and 1996 respectively. His PhD thesis received the 1996 Wilts Prize and the 1996 Clauser Prize at Caltech. From 1996 to 1997 he was a postdoctoral associate at MIT. Between 1997 and 2005 he was on the faculty the Electrical Engineering Department at UCLA, reaching the rank of Associate Professor. Since 2005 he is Professor of Electrical and Telecommunications Engineering at Universidad ORT Uruguay, and currently Vice-Dean of Research.

Dr. Paganini has received the 1995 O. Hugo Schuck Best Paper Award, the 1999 Packard Fellowship, the 2004 George S. Axelby Best Paper Award. He is a member of the Uruguayan National Academy of Sciences, the Uruguayan National Academy of Engineering, and the Latin American Academy of Sciences. During the pandemic he served in Uruguay as one of three coordinators of the Honorary Scientific Advisory Group on Covid-19, receiving after the Presidency of the Republic Award. He is a Fellow of the IEEE (2014) and a Fellow of IFAC (2023). His research interests are control and networks.

Andres Ferragut obtained his PhD in Electrical Engineering (Telecommunications) from Universidad de la Republica, Uruguay (2011). He is currently Full Professor of Networks and Communication Systems at Universidad ORT Uruguay, working in the Mathematical Analysis in Telecommunications and Energy (MATE) research group. In 2012, Dr. Ferragut was awarded the annual Uruguayan Engineering Academy prize for the best doctoral thesis in Electrical Engineering. He served as Editor for IEEE/ACM Transactions on Networking during the period 2016-2020. His research interests are stochastic processes and queueing theory applied to the mathematical modeling of networks.