An incentive algorithm for a closed stochastic network: data and mean-field analysis

Bianca Marin Moreno, Christine Fricker, Hanene Mohamed, Amaury Philippe, Martin Trépanier

An incentive algorithm for a closed stochastic network: data and mean-field analysis

Bianca Marin Moreno[1], Christine Fricker[1], Hanene Mohamed[2], Amaury Philippe[3] and Martin Trépanier[3]

[1]INRIA de Paris, DI-ENS, PSL Research University, Paris, France.
[2]MODAL'X, UMR CNRS 9023, UPL, Université Paris Nanterre, Paris, France.
[3]Mathematical and Industrial Engineering, Polytechnique Montreal, Montreal, H3T 1J4, Canada.

Contributing authors: bianca.marin-moreno@inria.fr; christine.fricker@inria.fr; hmohamed@parisnanterre.fr; amaury.philippe@polymtl.ca; martin.trepanier@polymtl.ca;

Abstract

The paper deals with a load-balancing algorithm for a closed stochastic network with two zones with different demands. The algorithm is motivated by an incentive algorithm for redistribution of cars in a large-scale car-sharing system. The service area is divided into two zones. When cars stay too long in the low-demand zone, users are encouraged to pick them up and return them in the high-demand zone. The zones are divided in cells called stations. The cars are the network customers. The mean-field limit solution of an ordinary differential equation (ODE) gives the large scale distribution of the station state in both clusters for this incentive policy in a discrete Markovian framework. An equilibrium point of this ODE is characterized via the invariant measure of a random walk in the quarter-plane. The proportion of empty and saturated stations measures how the system is balanced. Numerical experiments illustrate the impact of the incentive policy. Our study shows that the incentive policy helps when the high-demand zone observes a lack of cars but a saturation must be prevented especially when the high-demand zone is small.

Keywords: Large scale analysis, mean-field, car-sharing, incentive algorithm, stochastic network, cluster, load balancing, closed Jackson networks, product-form distribution

MSC Classification: 35A01 , 65L10 , 65L12 , 65L20 , 65L70

Acknowledgements

The authors would like to thank Communauto for providing data, funding and allowing to do this study. They also thank the Natural Science and Engineering Research Council of Canada (NSERC) for funding.

1 Introduction

Motivation. Car-sharing, a practice that is gaining ground in urban areas, comes to meet ecological, economic and practical imperatives. For a decade it has been becoming an alternative mode of transportation. The principle is that a given number of vehicles are made available to users at stations or in a public space in a given geographical area. The user picks up a vehicle if available, makes his trip and then drops it off at his destination.

For the operator, managing such systems is far from simple. The randomness due to the user arrivals as well as to the trips generates an imbalance in the system: some areas are more or less served by vehicles throughout the day, depending on whether they are residential areas or in the city center for example. Thus, the users may not find an available vehicle, which alters the efficiency of the system. A major issue for the operators is to balance the network by locating the vehicles more efficiently, i.e. getting them where they are needed. The usual techniques are either active, such as using trucks to move bikes or drivers for cars, or passive, such as incentive policies that encourage users to move vehicles themselves on their trips. We can cite the example of Velib+ which offered extra time for returning bikes in uphill stations of the Parisian bike-sharing system or the Angel's Rewards bikes program developed in NYC allowing users to earn free day passes and membership extensions.

Gift incentive policy. This paper deals with an incentive policy implemented by Communauto on its free floating car-sharing system in Montreal. In the geographical area, a small zone is identified as a high-demand zone by the operator. Some cars remain stationary for too long in the rest of the service area (with low-demand, called the normal zone) while users cannot find available cars in the high-demand zone. In order to bring back these stagnant cars from the normal zone to the high-demand zone, Communauto designates them as *gifts* on its smartphone app and offers 30 free minutes on the trip if the user returns the *gift* to the high-demand zone. This policy is called here the *gift policy*.

Aim of the paper. The aim is to study the impact of the incentive policy implementing a trip discount to move some cars to a high-demand area.

For this purpose, an analysis is done on a dataset provided by the operator in order to better understand the user behavior towards this incentive policy,

and more generally their use of the system. Based on this, a probabilistic model is proposed for such a system, as a large closed stochastic network of interacting particles which are cars and gifts. The service area is divided into cells, called here *stations*, which are nodes of the network, plus extra-nodes containing moving cars and gifts.

Results. We investigate in a Markovian framework the steady state of these stations. Although an invariant measure exists for this irreducible Markov process on a finite state space for a fixed number of stations, it has no explicit form as far as we know. The idea is to deal with the approximation as the number of stations and cars get large together, called the mean-field limit. Indeed, the states of the stations are asymptotically independent and their common distribution is given as a solution of an ordinary differential equation (ODE). See Proposition 1. The equilibrium point of the ODE gives the long-time limit. The special case of a model without incentive policy corresponds to the two-cluster model studied in [1] where the equilibrium point is unique and well determined. The two-cluster model of [1] finds in this paper a practical application through this car-sharing system with a service area with an intensive zone, but without an incentive policy. For the *gift policy*, Proposition 2 gives a characterization of the equilibrium point as a function of the invariant measure of a random walk in the quarter-plane. It is a first step to address the problem of existence and uniqueness of the equilibrium point.

Performance. Our performance criterion is to minimize the proportion of empty or saturated stations, called for short *problematic*, in order to maximize the efficiency of the system. Since no closed-form solution for the previous invariant measure is derived, we perform a numerical solution of a multidimensional equation for the system with incentive policy. We compare it with the analytical solution of the model without incentives. We study the impact of the policy in the case where everyone follows incentives. This impact is significant when the high-demand zone lacks cars. The risk is to overload it, especially if it is small.

Related works. The mean-field approach comes from physical statistics and is applied for a large class of models of interacting particles. For example biology (neuron networks), economics and social sciences (opinion dynamics). It has applications for stochastic networks as in the context of load-balancing strategies as Power-of-Choice and others. See references in [2] and [3]. A direct analysis of large networks is difficult because there is either no closed-form expression in general or some but numerically prohibitive. Thus the goal is to find approximations. The mean-field techniques provide an approximation of the evolution of the state distribution of a fixed node as a solution of an ODE. One of the main problems to address is the number of equilibrium points. This issue elucidated, it allows for the performance metrics as either explicit expressions or a numerical solution (see [4]), depending on the models. In all these models, the mean-field limit stands for the completely connected network while in real systems, the interaction is often weaker. A great difference

between the models is the difficulty to address the uniqueness of the equilibrium point. For example, in our study this question is not addressed since it is numerically solved. Proving existence of multiple equilibrium points can also be hard. In opinion or epidemic models, the set of states of an agent is small (even $\{0, 1\}$) which simplifies the analysis. See [5].

For a direct approach to stochastic network models, let us briefly recall an important class of queueing networks called *closed Jackson networks*. A closed network is a network where customers stay forever in the system with no external arrivals. A Gordon–Newell network, or simply called closed Jackson network [6], consists of a finite number of nodes, each with one or several servers. The service times of customers at each node are i.i.d. with exponential distribution. The service rate at each node can depend on both the node and its state. Specifically, if there are x_i customers at node i, the service rate is denoted by $\mu_i(x_i)$, where $\mu_i(.)$ is a function from $\mathbb{N} \to \mathbb{R}_+$ with $\mu_i(0) = 0$ and $\mu_i(x) > 0$ for all $x > 0$. At the end of its service, a customer is directed to another node according to a fixed routing matrix. All service times are independent. Note that a Jackson network is the version with external arrivals and a routing matrix including departures to the outside (see [7]). Originally, the nodes of such a network are assumed to have infinite capacity. This description can be generalized to a system including nodes with finite capacities. Nevertheless one should describe the way customers must be redirected until reaching non-saturated nodes, i.e. whose capacity is not yet reached. Consider the blocking-rerouting policy where entering a saturated queue, the customer is rerouted with the routing matrix at infinite speed until he finds an unsaturated node. This blocking-rerouting policy is defined in [8] and used in [9] in a more general version. The main interest of such a class of queueing networks is that the equilibrium distribution is explicit with a product-form expression (see [8]).

It is conventional to model vehicle-sharing systems as closed Jackson networks where the vehicles play the role of the customers. The associated network has two types of nodes: one-server nodes, that describe the stations, and infinite-server nodes, the latter corresponding to the different routes linking the stations. Service times at these nodes are respectively inter-arrival times of users at stations and trip times along the corresponding routes. In a pioneering paper [10], an asymptotic analysis of infinite capacity closed Jackson networks at equilibrium is proposed when the number of nodes and customers tend to $+\infty$ at the same rate. It is applied to vehicle-sharing systems in the infinite capacity case. Note that the asymptotic analysis performed in [11] is done when the number of nodes (stations and routes) is fixed, while the number of customers (vehicles) tends to infinity. Both papers crucially rely on the explicit product-form stationary distribution, which is well-known in the infinite capacity case. The case of vehicle-sharing systems with finite capacity is considered in [9]. The model is identified as a Jackson network with the blocking-rerouting policy previously described. By [8], the invariant distribution has a product-form in this case.

Outline of the paper. In Section 2, an analysis is carried out on the data provided by the Montreal operator for the free-floating car-sharing system, which motivates the definition of our model. Section 3 deals with the model description, the notations and the Markov process of the state of the stations with its $\mathcal{Q}$-matrix. In Section 4, the main result giving the mean-field limit as a solution of an ODE is stated in Proposition 1. In Section 5, the equilibrium point of the ODE is investigated in the model without gifts and with gifts. Section 6 compares the performance in both models. A discussion focused on the model assumptions is proposed in Section 7. Section 8 is devoted to future work.

2 Data Analysis

2.1 Data Set

The dataset analyzed in this paper comes from Communauto, the unique car-sharing operator in Montreal, Canada. It concerns its free-floating car-sharing system serving Montreal and covers the period from October 2019 to October 2020. In order to avoid the effects of the first lockdown of Covid-19, the three months from March to May 2020 are not taken into account. Data is also divided in Winter (October 2019 - February 2020) and Summer (June 2020 - October 2020) when analyzing the influence of the seasons.

The dataset consists of 1,480,473 transactions of which 52.2% was in winter and 47.8% in summer. The service area as shown on Communauto's app covers nearly 100 km^2 (see [12]), and is divided into a high-demand zone, called zone 1 in the following, with an estimated area of 28 km^2 that can be seen in Figure 1 and a low-demand zone called zone 2 with an area of approximately 72 km^2 (refer to Communauto's app). The GPS positions of the boundaries of the high-demand zone are used to determine in the dataset which zone a car is in. The features available to us within the data to perform this analysis are, for each transaction,

- Date and time of the start and end of the transaction
- Latitude and longitude of the start and end of the transaction
- Total fate (the sum of the price due to the duration of the transaction and the price due to the distance of the transaction)
- Identifier of the vehicle and of the user
- Distance traveled.

To determine whether a transaction has started or ended in zone 1 or not, we use Communauto's application map. We approximate the high-demand zone by a polygon and compute the latitude and longitude of its vertices. By making a change of coordinates from spherical to planar, we write an algorithm to decide if a point is inside the polygon or not.

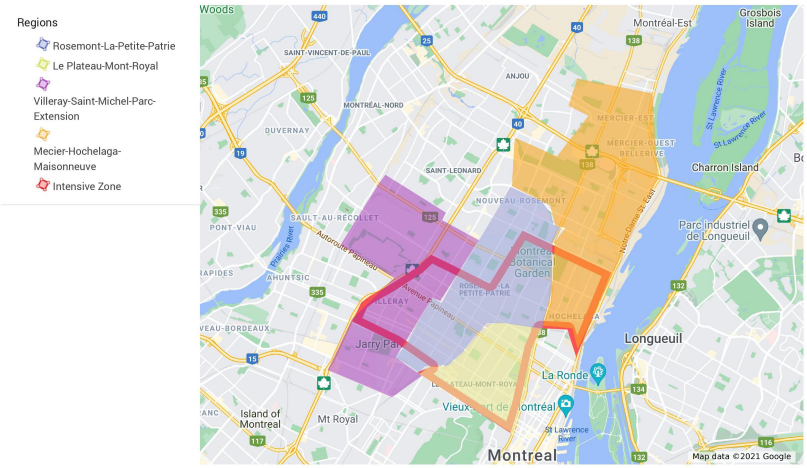

Fig. 1: Communauto Free-floating Car-Sharing service area and its high-demand zone circled in red. Source: Communauto's app on 2021/05/10.

2.2 Classifying *Gifts* within the data

As explained in the introduction, this paper is interested in studying an incentive policy in the free floating system of Montreal that we call the *gift incentive policy*. To do so, we need to define in the data which transactions are performed by cars that appear in Communauto's application with a gift symbol, which we call *gifts* for short, and the *gifts* transactions that contribute to the redistribution objective of the policy, which we call *used gifts*.

A transaction is classified as a *gift* if it satisfies two criteria:

- the car used for the transaction has been *stationary* for more than 16 hours since its last transaction,
- the car used for the transaction is parked in the low-demand zone (zone 2).

In the first criterion, being *stationary* means that the car has not moved, which may include the case where the user makes a reservation and then cancels it. A previous study based on a dataset of the Montreal free floating system from 2014 to 2017 indicates that the average reservation lasts 22 minutes [13]. Since this time is negligible compared to the 16 hours that a car must be *stationary* to be a *gift*, we conclude that reservation does not have too much impact on the gift policy and thus is not considered in this study.

A transaction is classified as a *used gift* if it is a *gift* and the car is returned in the high-demand zone (zone 1).

Table 1: Proportion and average daily number of transactions that are eligible as a *gift* due to remaining *stationary* for over 16 hours (*gift* by time), that are *gifts*, and that are *used gifts*. Winter is from October 2019 to February 2020 and Summer is from June 2020 to October 2020.

Period	Classification	Proportion (%)	Average daily number
Winter	Eligible as *gift* by time	1.69	58.5
	gift	1.36	47
	used gifts	0.42	14.5
Summer	Eligible as *gift* by time	3.56	114
	gift	2.67	86
	used gifts	0.92	29.5

2.3 Impact of the *gift policy*

Having detected the transactions that are *gifts* and *used gifts* we can now analyse the impact of the *gift policy* in the Montreal free-floating system. Such analysis assists to build a relevant model.

2.3.1 Gifts within data

Table 1 gives the proportion of *gifts* and *used gifts* detected among the transactions according to the seasons. It seems that in the summer, probably due to the vacation period, the cars stay idle longer than during winter. The proportion of *gifts* and *used gifts* among the transactions is very low. On a day, in average 0.42% are gifts in winter and 0.92% in summer.

Figure 2 plots the neperian logarithm of the frequency of transactions made by cars that stay *stationary* for at least x hours. This curve appears to follow a Zipf's law (see [14]) with a shape parameter of 1.65. Furthermore, we find that 80% of the cars are used before remaining *stationary* for over 5 hours.

2.3.2 Popularity

Over 90% of Communauto users have never made a *gift* transaction. Moreover, we find that few users are responsible for the majority of *used gift* transactions and that most users who have ever done a *used gift* transaction have done so only once. For each user in the dataset, we count the number of *used gifts* transactions and rank them according to their frequency of use. The top-ranked user is the most frequent one being responsible for 2.6% of all *used gifts* trips. The 10th user is responsible for 0.6% of *used gift* transactions. Finally, 70% percent of users among those with at least one *used gift* transaction make only one such transaction and are responsible for 33% of all *used gift* transactions. Moreover, Figure 3 shows that the frequencies as a function of rank in a logarithm scale is a straight line, following a Zipf distribution with parameter 0.6.

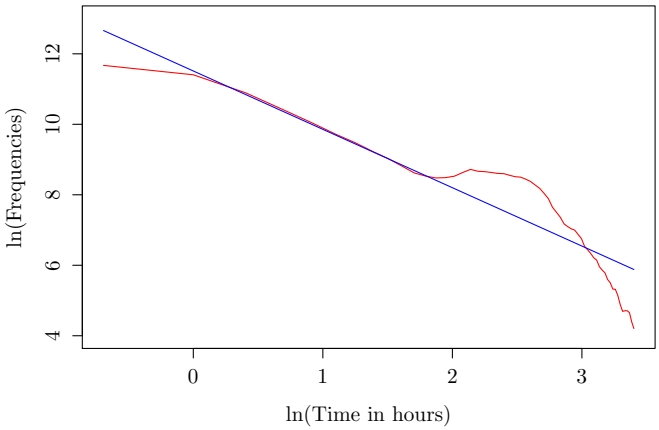

Fig. 2: Transaction frequency as a function of the time the car is *stationary* before the transaction. The log-log plot shows a Zipf distribution with a shape parameter of 1.65.

2.3.3 A redistribution during the day

In Figure 4, the green curve with the axis on the right represents the daily evolution of the number of *gifts* available at each hour of the day. The peak, of 25 gifts, is reached around noon. Thus, a user sees no more than 25 gifts out of 500 cars on the application at a time. A comparison can be made with the red and blue curves in the same figure (left axis), representing the average daily number of cars parked in each zone. The red curve for the high-demand zone reaches its minimum at 1pm, one hour after the gift peak. Hence, we conclude that the redistribution of cars from the low-demand zone to the high-demand zone by the *gift policy* is the strongest exactly when the number of cars available in the high-demand zone is the lowest, implying that this policy redistributes cars when they are most needed. Moreover, we see that it redistributes during the day, which is expected for an incentive policy at a time when using corporate drivers would make some cars unavailable to users.

2.3.4 Looking for gift price discounts on data

We can ask how much someone saves by taking a *gift* to make a transaction between the low-demand zone to the high-demand zone instead of taking a normal car. The final price of a transaction is composed of a distance rate plus a time rate. According to Communauto's *gift policy*, when users return *gifts* in the intensive zone, they earn 30 free minutes. From the dataset, the average travel time for *used gifts* is of 1.24 hours, and for other transactions from the normal zone to the intensive zone is of 1.13 hours. On the other hand,

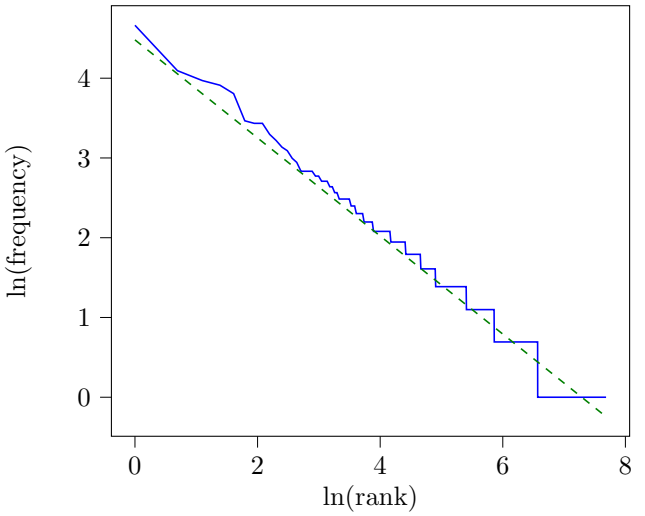

Fig. 3: Zipf's law with shape parameter 0.6 (frequency a *gift* user uses a *gift*)

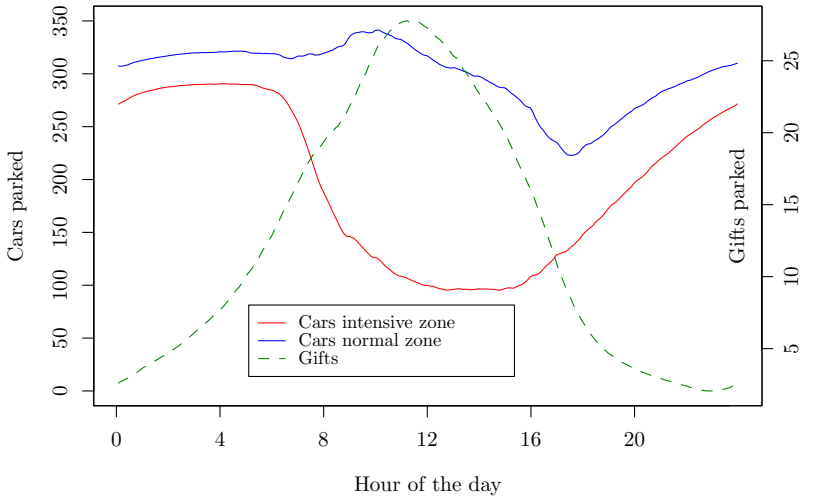

Fig. 4: Changes in the average daily number of *gifts* (right axis) and cars (left axis) parked using a 10-month data set (10/2019-10/2020 except 03-05/2020). The time steps are of 5 minutes.

the average time rate for *used gifts* is 4.62 dollars, while the average time rate for normal cars is 6.14 dollars. Thus, using a *gift* for the transaction between zone 2 to 1 is 31.3% times cheaper than the average time rate for normal transactions with the same trajectory.

2.4 Data analysis conclusion

This data analysis allows us to better understand how Communauto's free floating system and *gift policy* work. We infer that *gifts* are used to regulate the system by redistributing vehicles exactly when needed. Furthermore, we see that the 30 free minutes incentive a user earns with a *used gift* indeed decreases the transaction price due to the transaction's time. However, less than 1% of transactions are *used gifts*. This makes it particularly difficult to detect changes in user behavior. In addition, 90% of Communauto users have never made a *gift* transaction, and nearly 70% of gift users have used a gift only once during the 10 months we analyse. Thus, few Communauto customers actively use the *gift policy* to contribute to the redistribution of vehicles.

3 The model

3.1 Model description

In this following description and in the whole paper, a car is always a *normal car* and not a gift. We propose a simplified stochastic two-cluster model for car-sharing systems when including the gift policy. It will be further discussed in Section 7.1. The model is described as follows.

- The arrival process of users at any station of cluster i is a Poisson process with parameter λ_i, where $i \in \{1, 2\}$. As the rate of user arrivals is larger in cluster 1 than cluster 2, we assume that $\lambda_1 > \lambda_2$.
- If the user arrives at a station in cluster 1 where there is available cars at this station, the user picks up one of them to start a trip. Otherwise he leaves the system.
- Every car parked in cluster 2 becomes a *gift* after a random time with exponential distribution of parameter δ.
- When a user arrives in a station of cluster 2, if there is an available *gift* and an available car in this station, he picks up a gift with probability p, and a car with probability $1 - p$. If there is just one of the resources (gift or car), the user picks it up. Otherwise he leaves the system.
- The car trip duration is assumed to have an exponential distribution of parameter μ. When a car trip ends, the user chooses cluster i with probability c_i, then he chooses a station at random in this cluster to park the car.
- The *gift* trip is assumed to have an exponential distribution of parameter μ_c. When a *gift* trip ends, the user returns the *gift* car to any station in cluster i with probability q_i. The *gift* parked appears then as a car on the app.
- A station in cluster i has capacity K_i. If the station chosen is full, the user makes another trip until finding a station with an available parking space.

Note that in our model the inter-arrival times of users, trip times and times to become a gift are all independent with exponential distribution. See Figure 5 for an illustration of the model. This modeling was preceded by an analysis of real data. The model will be discussed in Section 7.1.

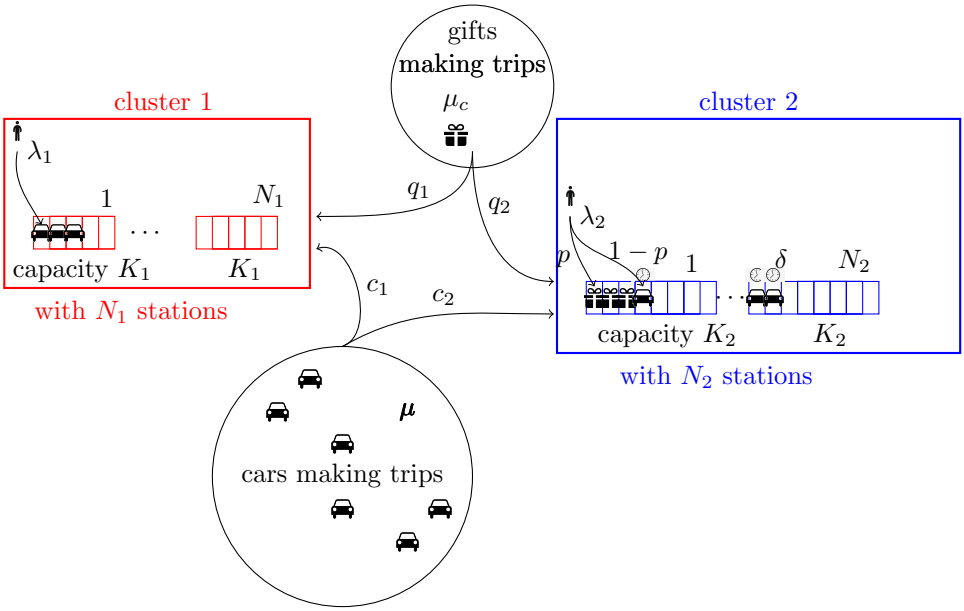

Fig. 5: Illustration of the model with gifts.

3.2 Notations

Let us summarize the notations. For all the following, $i \in \{1, 2\}$ is the cluster type.

- N_i is the number of stations in cluster i.
- $N = \sum_i N_i$ is the total number of stations.
- $\alpha_i = \lim_{N\to+\infty} N_i/N$ is the limiting proportion of stations in cluster i.
- K_i is the capacity of a station in cluster i.
- M is the total number of cars.
- $s = \lim_{N\to+\infty} M/N$ is the limiting mean number of cars per station, called *fleet size parameter*.
- λ_i is the rate of user arrivals at a station in cluster i.
- $1/\mu$ is the mean trip time for a normal car.
- $1/\mu_c$ is the mean trip time for a *gift*.
- δ is the rate at which a car in a station of cluster 2 becomes a *gift*.
- p is the probability that a user takes a gift when cars and gifts are both available.

- q_i is the probability that a user returns the gift in cluster i.
- c_i is the probability that a user returns his normal car in cluster i.

3.3 Queueing formulation

In this paper, the system can be described as a closed stochastic network. The nodes of the network are a set of $N = N_1 + N_2$ one-server queues of finite capacity (the stations), divided in two clusters, cluster 1 (the zone with high demand) with N_1 stations of capacity K_1, cluster 2 (the normal zone) with N_2 stations of capacity K_2, plus two infinite-server queues, i.e. the two nodes containing respectively cars and gifts making a trip. The service times at the queues have exponential distribution with parameters respectively λ_1, λ_2, μ and μ_c. According to the queueing vocabulary, there are M customers of two classes: cars and gifts, and a routing matrix given by the previous description.

However, this is not a Jackson network because there are additional transitions since a car in a station of cluster 2 becomes a gift at rate δ and a gift arriving at a station from the infinite-server node becomes a car. It does not fit in this classical framework because of these changes of customer classes. Note that in the case without incentive policy ($\delta \to 0$), the model is a Jackson network since there are no gifts. Such a model (with $\delta \to 0$) is known as a two-cluster bike-sharing system studied in [1] and [9]. Section 5.1 is devoted to this case called *model without gifts*.

3.4 The Markov process

The state process is

$$\big(X_{1,n}(t), X_{2,m}(t), C_m(t), Z^N(t),\ 1 \le n \le N_1 \text{ and } 1 \le m \le N_2\big)$$

where

- $X_{1,n}(t)$ is the number of cars at a station n in cluster 1 at time t,
- $X_{2,m}(t)$ is the number of cars at a station m in cluster 2 at time t,
- $C_m(t)$ is the number of *gifts* at a station m (necessarily in cluster 2) at time t and
- $Z^N(t)$ is the number of *gifts* making a trip at time t.

Note that the number of cars making a trip at time t is equal to

$$M - \sum_{n=1}^{N_1} X_{1,n}(t) - \sum_{m=1}^{N_2} \big(X_{2,m}(t) + C_m(t)\big) - Z^N(t).$$

As we deal with a two-cluster model, it is sufficient to study the behavior of one station in each cluster. It amounts to dealing with the empirical measure

process

$$(Y^N(t)) = \Big(Y^{N_1}_{1,j}(t), Y^{N_2}_{2,k,l}(t), \frac{Z^N(t)}{N}, j \in \chi_1, (k,l) \in \chi_2\Big)$$

where $Y^{N_1}_{1,j}(t)$ is the proportion of stations with j cars in cluster 1 and $Y^{N_2}_{2,k,l}(t)$ is the proportion of stations with k cars and l *gifts* in cluster 2, defined by

$$Y^{N_1}_{1,j}(t) = \frac{1}{N_1}\sum_{n=1}^{N_1} \mathbf{1}_{\left\{X_{1,n}(t)=j\right\}} \text{ and } Y^{N_2}_{2,k,l}(t) = \frac{1}{N_2}\sum_{m=1}^{N_2} \mathbf{1}_{\left\{\left(X_{2,m}(t),C_m(t)\right)=(k,l)\right\}}$$

where $\chi_1 = \{j \in \mathbb{N},\ j \le K_1\}$ and $\chi_2 = \{(k,l) \in \mathbb{N}^2,\ k+l \le K_2\}$. Because the inter-arrival times, trip times and times to become a gift have exponential distribution, $(Y^N(t))$ is a Markov process, with finite state space

$$S^N = \Big\{ y = (y_{1,j}, y_{2,k,l}, z)_{\{j\in\chi_1,(k,l)\in\chi_2\}}, y_{1,j} \in \frac{\mathbb{N}}{N_1}, y_{2,k,l} \in \frac{\mathbb{N}}{N_2}, z \in \frac{\mathbb{N}}{N},$$

$$\sum_{j\in\chi_1} y_{1,j} = 1, \sum_{(k,l)\in\chi_2} y_{2,k,l} = 1, \sum_{j\in\chi_1} j\, y_{1,j}\, N_1 + \sum_{(k,l)\in\chi_2} (k+l)\, y_{2,k,l}\, N_2 + z\, N \le M \Big\}.$$

The inequality in the previous definition of the state space S^N is due to the fact that the number of cars driving has be to added to the left-hand side of the inequality to obtain the total number M of cars in the system. Let us write its transitions from state $y \in S^N$. To simplify the notations, let us denote by

$$E_1 = \sum_{j\in\chi_1} j\, y_{1,j} \quad \text{and} \quad E_2 = \sum_{(k,l)\in\chi_2} (k+l)\, y_{2,k,l} \tag{1}$$

respectively the mean number of cars parked per station in cluster 1 and the mean number of cars plus gifts parked per station in cluster 2. Also, let us denote by $(e_{1,j}, e_{2,k,l}, e_0, j \in \chi_1,\ (k,l) \in \chi_2)$ the canonical basis of $\mathbb{R}^{|\chi_1|+|\chi_2|+1}$, where the cardinality of set A is denoted by $|A|$. The transitions, from state $y = (y_{1,j}, y_{2,k,l}, z) \in S^N$, are due to three events: a user arrival, a gift appearance or a car return. The transitions, from state $y = (y_{1,j}, y_{2,k,l}, z) \in S^N$, are given by the non-null non-diagonal terms of $\mathcal{Q}$-matrix Q^N defined as follows.

- **User arrival.**
  - A user arrival at a station in cluster 2 with k cars and l gifts (for short of type $(2,k,l)$) taking a gift. The number of gifts decreases by 1. This happens if $l > 0$. There are $y_{2,k,l}N_2$ possible stations. Thus this event happens at rate $\lambda_2 y_{2,k,l} N_2 \mathbf{1}_{\{l>0\}}\big(p + (1-p)\mathbf{1}_{\{k=0\}}\big)$. Recall that p is the probability for a user arriving to a station in cluster 2 to choose a gift

when cars and gifts are available.

$$Q^N(y, y + \frac{1}{N_2}(e_{2,k,l-1} - e_{2,k,l}) + \frac{e_0}{N}) = \lambda_2 y_{2,k,l} N_2 \mathbf{1}_{\{l>0\}} \big(p + (1-p)\mathbf{1}_{\{k=0\}}\big).$$

– A user arrival at a station of type $(2, k, l)$ taking a normal car.

$$Q^N\left(y, y + \frac{1}{N_2}(e_{2,k-1,l} - e_{2,k,l})\right) = \lambda_2 y_{2,k,l} N_2 \mathbf{1}_{\{k>0\}} \left(1 - p + p\mathbf{1}_{\{l=0\}}\right).$$

– A user arrival at a station of type $(1, j)$.

$$Q^N\left(y, y + \frac{1}{N_1}(e_{1,j-1} - e_{1,j})\right) = \lambda_1 y_{1,j} N_1 \mathbf{1}_{\{j>0\}}.$$

- **Gift appearance.**
  – A car becoming a gift at a station of type $(2, k, l)$.

$$Q^N\left(y, y + \frac{1}{N_2}(e_{2,k-1,l+1} - e_{2,k,l})\right) = \delta k N_2 y_{2,k,l}.$$

- **Car return.**
  – A normal car returned at a station of type $(1, j)$.

$$Q^N\left(y, y + \frac{1}{N_1}(e_{1,j+1} - e_{1,j})\right) = c_1 y_{1,j} \mu \left(M - E_1 N_1 - E_2 N_2 - zN\right) \mathbf{1}_{\{j<K_1\}}.$$

  – A normal car returned at a station of type $(2, k, l)$.

$$Q^N\left(y, y + \frac{1}{N_2}(e_{2,k+1,l} - e_{2,k,l})\right) = c_2 y_{2,k,l} \mu \left(M - E_1 N_1 - E_2 N_2 - zN\right) \mathbf{1}_{\{k+l<K_2\}}.$$

  – A gift returned at a station of type $(1, j)$.

$$Q^N\left(y, y + \frac{1}{N_1}(e_{1,j+1} - e_{1,j}) - \frac{e_0}{N}\right) = q_1 y_{1,j} \mu_c z N \mathbf{1}_{\{j<K_1\}}.$$

  – A gift returned at a station of type $(2, k, l)$.

$$Q^N\left(y, y + \frac{1}{N_2}(e_{2,k+1,l} - e_{2,k,l}) - \frac{e_0}{N}\right) = q_2 y_{2,k,l} \mu_c z N \mathbf{1}_{\{k+l<K_2\}}.$$

These transitions allow us to write the drift of process $(Y^N(t))_t$ which will be useful to the mean-field convergence (Proposition 1).

4 Mean-field limit

Our aim is to investigate the system when M, N_1 and N_2 get large at the same rate, for short, when N gets large. When N tends to $+\infty$, the process $(Y^N(t))$ given by the previous transitions converges in distribution to a deterministic function which is the unique solution of a given ODE. This result is given by the following proposition.

Proposition 1 (Mean-field convergence) *For $T > 0$, $(Y^N(t))_{t\in[0,T]}$ converges in distribution to the unique solution $(y(t))_{t\in[0,T]}$ of the following ODE with $y(0)$ fixed*

$$
\begin{aligned}
&\frac{\mathbf{dy_{1,j}}}{\mathbf{dt}}(\mathbf{t}) = y_{1,j+1}(t)\lambda_1 \mathbf{1}_{\{j<K_1\}} \\
&+ y_{1,j-1}(t)\left(\frac{c_1}{\alpha_1}\mu(s-\alpha_1 E_1(t)-\alpha_2 E_2(t)-z(t)) + \frac{q_1\mu_c}{\alpha_1}z(t)\right)\mathbf{1}_{\{j>0\}} \\
&- y_{1,j}(t)\left(\lambda_1 \mathbf{1}_{\{j>0\}} + \frac{c_1}{\alpha_1}\mu(s-\alpha_1 E_1(t)-\alpha_2 E_2(t)-z(t))\mathbf{1}_{\{j<K_1\}} + \frac{q_1\mu_c}{\alpha_1}z(t)\mathbf{1}_{\{j<K_1\}}\right) \\
&\frac{\mathbf{dy_{2,k,l}}}{\mathbf{dt}}(t) = y_{2,k,l+1}(t)\lambda_2 \mathbf{1}_{\{k+l<K_2\}}\big(p+(1-p)\mathbf{1}_{\{k=0\}}\big) \\
&+ y_{2,k+1,l}(t)\lambda_2 \mathbf{1}_{\{k+l<K_2\}}\big(1-p+p\mathbf{1}_{\{l=0\}}\big) + y_{2,k+1,l-1}(t)\,\delta\,(k+1)\,\mathbf{1}_{\{k<K_2\}} \\
&+ y_{2,k-1,l}(t)\left(\frac{c_2}{\alpha_2}\,\mu\,(s-\alpha_1 E_1(t)-\alpha_2 E_2(t)-z(t)) + \frac{q_2\mu_c}{\alpha_2}z(t)\right)\mathbf{1}_{\{k>0\}} \\
&- y_{2,k,l}(t)\Bigg(\lambda_2(1-\mathbf{1}_{\{k=0,l=0\}}) + \delta k + \frac{c_2}{\alpha_2}\mu(s-\alpha_1 E_1(t)-\alpha_2 E_2(t) \\
&- z(t))\mathbf{1}_{\{k+l<K_2\}} + \frac{(q_2\mu_c}{\alpha_2}z(t)\mathbf{1}_{\{k+l<K_2\}}\Bigg) \\
&\frac{\mathbf{dz}}{\mathbf{dt}}(\mathbf{t}) = -q_1\,\mu_c z(t)\sum_{j\in\chi_1} y_{1,j}(t)\,\mathbf{1}_{\{j<K_1\}} + \alpha_2\,\lambda_2 \sum_{(k,l)\in\chi_2} y_{2,k,l}(t)\,\mathbf{1}_{\{l>0\}}\,\big(p+(1-p)\mathbf{1}_{k=0}\big) \\
&- q_2\mu_c z(t)\sum_{(k,l)\in\chi_2} y_{2,k,l}(t)\mathbf{1}_{\{k+l<K_2\}}. \qquad (2)
\end{aligned}
$$

Recall that, in these equations, s is the limiting number of cars per station and α_i the limiting proportion of stations in cluster i, $i \in \{1,2\}$.

Proof The proof is standard (see [15]). The idea of the proof is that a Markov process can be written as the sum of a martingale term and a drift term in form of an integral on time. When N is large, one can prove that the process is tight. Moreover, the martingale term converges to 0. Then any limiting value satisfies an ODE. The uniqueness of the solution of the ODE gives the convergence of the process. Let us present a sketch of the proof.

The key argument is to write the evolution equations of process $(Y^N(t))$. It is omitted here. It gives that

$$Y^N(t) = Y^N(0) + \int_0^t Q^N(id)(Y^N(s))ds + M^N(t)$$

where $M^N(t)$ is a martingale. Then, using the expression of Q-matrix Q^N, the previous equation can be rewritten

$$
\begin{aligned}
Y^N(t) = Y^N(0) + \sum_{(k,l)\in\chi_2} \int_0^t Y^N_{2,k,l}(s) \\
\Big((p + (1-p)\mathbf{1}_{\{k=0\}})\lambda_2 \mathbf{1}_{\{l>0\}} \left(e_{2,k,l-1} - e_{2,k,l} + \frac{e_0 N_2}{N} \right) \\
+ \lambda_2 \mathbf{1}_{\{k>0\}} \left(1 - p + p\mathbf{1}_{\{l=0\}} \right) (e_{2,k-1,l} - e_{2,k,l}) \\
+ \frac{c_2\mu}{N_2} \left(M - E_1^N(s)N_1 - E_2^N(s)N_2 - Z^N(s)N \right) \mathbf{1}_{\{k+l<K_2\}} (e_{2,k+1,l} - e_{2,k,l}) \\
+ \delta k (e_{2,k-1,l+1} - e_{2,k,l}) + q_2\mu_c Z^N(s) N \mathbf{1}_{\{k+l<K_2\}} \left(\frac{1}{N_2}(e_{2,k+1,l} - e_{2,k,l}) - \frac{e_0}{N} \right) \Big) ds \\
+ \sum_{j\in\chi_1} \int_0^t Y^N_{1,j}(s) \\
\Big(\lambda_1 \mathbf{1}_{\{j>0\}} (e_{1,j-1} - e_{1,j}) - e_0 q_1 \mu_c Z^N(s) \mathbf{1}_{\{j<K_1\}} \\
+ \left(c_1 \frac{\mu}{N_1} \left(M - E_1^N(s)N_1 - E_2^N(s)N_2 - Z^N(s)N \right) + q_1\mu_c Z^N(s) \frac{N}{N_1} \right) \\
\mathbf{1}_{\{j<K_1\}} (e_{1,j+1} - e_{1,j}) \Big) ds + M^N(t)
\end{aligned}
$$

where $E_i^N(t)$, $i = 1, 2$ are defined from E_i given by equation (1) replacing y by $Y^N(t)$. Using the evolution equation, one can prove the tightness of $(Y^N(t))$ and that the martingale term $(M^N(t))$ tends to 0. Plugging in the previous equation, any limiting point $(y(t))$ of $(Y^N(t))$ satisfies

$$
\begin{aligned}
\frac{dy(t)}{dt} = \sum_{(k,l)\in\chi_2} y_{2,k,l}(t) \Big((p + (1-p)\mathbf{1}_{\{k=0\}})\lambda_2 \mathbf{1}_{\{l>0\}} \left(e_{2,k,l-1} - e_{2,k,l} + e_0\alpha_2 \right) \\
+ \lambda_2 \mathbf{1}_{\{k>0\}} \left(1 - p + p\mathbf{1}_{\{l=0\}} \right) (e_{2,k-1,l} - e_{2,k,l}) \\
+ \frac{c_2\mu}{\alpha_2} \left(s - \alpha_1 E_1(t) - \alpha_2 E_2(t) - z(t) \right) \mathbf{1}_{\{k+l<K_2\}} (e_{2,k+1,l} - e_{2,k,l}) \\
+ \delta k (e_{2,k-1,l+1} - e_{2,k,l}) + q_2\mu_c z(t) \mathbf{1}_{\{k+l<K_2\}} \left(\frac{1}{\alpha_2}(e_{2,k+1,l} - e_{2,k,l}) - e_0 \right) \Big) \\
+ \sum_{j\in\chi_1} y_{1,j}(t) \Big(\lambda_1 \mathbf{1}_{\{j>0\}} (e_{1,j-1} - e_{1,j}) - e_0 q_1 \mu_c z(t) \mathbf{1}_{\{j<K_1\}} \\
+ \left(\frac{c_1\mu}{\alpha_1} \left(s - \alpha_1 E_1(t) - E_2(t)\alpha_2 - z(t) \right) + q_1 \frac{\mu_c z(t)}{\alpha_1} \right) \mathbf{1}_{\{j<K_1\}} (e_{1,j+1} - e_{1,j}) \Big)
\end{aligned}
$$

which is equivalent to equation (2). But this ODE has a unique solution with fixed $y(0)$. This ends the proof. □

5 The equilibrium point

To investigate the steady-state behavior of the model, we study the equilibrium point $\bar{y}$ of the mean-field ODE written as follows

$$\frac{\mathbf{dy}}{\mathbf{dt}}(\mathbf{t}) = F(y(t))$$

where F comes from Proposition 1. It amounts to finding $\bar{y}$ such that

$$F(\bar{y}) = 0. \tag{3}$$

Note that the vector $\bar{y}$ is of dimension $1 + |\chi_1| + |\chi_2| = 1 + K_1 + K_2(1 + K_2)/2$. Finding a closed-form expression of the equilibrium point $\bar{y}$ is out of reach. Let us present two points of view: the first one is based on a nice queueing interpretation which holds for the no-gift case. The second is an analytic approach which should be relevant for the case with gifts but is beyond this work.

5.1 The queueing interpretation for the no-gift case

In this case, the existence and uniqueness of the equilibrium point $\bar{y}$ is proved. See [1] for details. In addition, $\bar{y}$ is given by a simple queueing interpretation of the mean-field limit. It gives that the limiting stationary number of cars at a station of cluster i, considered as a $M/M/1/K_i$ queue, has a geometric distribution $\nu_{\rho r_i, K_i}$ on $\{0, \dots, K_i\}$ with parameter ρr_i where for $i = 1, 2$, $r_i = \Lambda\mu\beta_i/\lambda_i$ with $\beta_i = q_i/\alpha_i$, $\Lambda = 1/\max_i(\mu\beta_i/\lambda_i)$ and ρ is the unique solution of the fixed point equation

$$s = \rho\Lambda + \sum_{i=1}^{2} \alpha_i m(\nu_{\rho r_i, K_i}). \tag{4}$$

In the previous equation, we denote by $m(\nu_{\rho,K})$ the mean of the geometric distribution $\nu_{\rho,K}$ on $\{0, \dots, K\}$ with parameter ρ, given by

$$m(\nu_{\rho,K}) = \begin{cases} \frac{K}{2} & \text{if } \rho = 1 \\ \frac{\rho}{1-\rho} - \frac{(K+1)\rho^{K+1}}{1-\rho^{K+1}} & \text{otherwise} \end{cases} \tag{5}$$

because, for $\rho = 1$, $\nu_{\rho,K}$ is the uniform distribution on $\{0, \dots, K\}$. It shows that the multidimensional equilibrium point equation (3) amounts to a fixed point equation (4) on $\mathbb{R}_+$. This is the purpose of [1, Theorem 1] for the cluster case detailed in [1, Section 2.3].

5.2 Characterization of the equilibrium point

Taking into account the gift policy induces a change of classes between normal cars and gifts. This considerably complicates the search for an equilibrium

point and changes the nature of the limiting objects involved. The question of existence and uniqueness of a solution of the equilibrium point in equation (3) remains open. For simplicity, let us take the case $p = q_1 = 1$ in order to highlight the main difficulties of this problem. Remembering that $p = 1$ means that, when available, a gift is always chosen over a car in a station of the normal zone, and $q_1 = 1$ means that all gifts are returned at a station of cluster 1. Heuristically, looking for an equilibrium point $\bar{y}$ means that the right-hand term in the mean-field ODE (2) is null. With obvious notations $\bar{y} = (\bar{y_1}, \bar{y_2}, \bar{z})$, note first that the number of moving gifts $(Z^N(t)/N)$ is the rescaled number of customers in a $M/M/\infty$ queue introduced by Kelly [16] whose limit is $(z(t))$ which tends to the load parameter (see [17, Section 6.5])

$$\bar{z} = \frac{\alpha_2 \lambda_2}{\mu_c} \frac{1 - \bar{y}_{2,.,0}}{1 - \bar{y}_{1,S}} \tag{6}$$

where $\bar{y}_{1,S}$ is the probability that a station in cluster 1 is saturated and $\bar{y}_{2,.,0}$ the probability that a station in cluster 2 has no gift, i.e. $1 - \bar{y}_{1,S} = \sum_{j \in \chi_1} \bar{y}_{1,j} \mathbf{1}_{\{j < K_1\}}$ and $1 - \bar{y}_{2,.,0} = \sum_{(k,l) \in \chi_2} \bar{y}_{2,k,l} \mathbf{1}_{\{l > 0\}}$. Then a queuing interpretation similar to that for the no-gift case holds. Indeed, at equilibrium, a station of cluster 1 can be considered as a $M/M/1/K_1$ queue, with arrival rate

$$\bar{\gamma_1} = \frac{1}{\alpha_1} \left(c_1 \mu (s - \alpha_1 \bar{E}_1 - \alpha_2 \bar{E}_2 - \bar{z}) + q_1 \mu_c \bar{z} \right) \tag{7}$$

where $\bar{E}_i$ are defined by (1) and service rate λ_1. It is well known that its invariant measure is a geometric distribution on $\{0, \ldots, K_1\}$ with parameter $\bar{\rho}_1 = \bar{\gamma}_1 / \lambda_1$, i.e. $\bar{y}_{1,j} = \rho_1^j (1 - \rho_1)/(1 - \rho_1^{K+1})$ for $0 \leq j \leq K_1$. Note that, plugging equation (6) into (7), $\bar{\rho}_1$ depends on $\bar{y}$, only by $\bar{y}_1$ and $\bar{y}_2$. Moreover $\bar{y}_2 = \pi_{\bar{\rho}_2, K_2}$ where

$$\bar{\rho}_2 = \frac{1}{\lambda_2 \alpha_2} \left(c_2 \mu (s - \alpha_1 \bar{E}_1 - \alpha_2 \bar{E}_2 - \bar{z}) + q_2 \mu_c \bar{z} \right)$$

and, for fixed ρ, π_{ρ, K_2} is the invariant measure of the Markov process on χ_2 with matrix jump $Q_{\rho,K}$ given by its non-null non-diagonal terms

$$\begin{cases} Q_{\rho,K}(n, n - e_1) & = \lambda_2 \mathbf{1}_{\{n_1 > 0\}} \\ Q_{\rho,K}(n, n + e_2) & = \lambda_2 \rho \mathbf{1}_{\{n_2 < K_2\}} \\ Q_{\rho,K}(n, n + e_1 - e_2) & = \delta \mathbf{1}_{\{n_2 > 0\}}. \end{cases} \tag{8}$$

In conclusion the equilibrium point $\bar{y}$, solution of a multidimensional fixed point equation, can be expressed as a function of $(\bar{\rho}_1, \bar{\rho}_2)$ solution of a fixed point equation. It is summarized by the following result.

Proposition 2 (Equilibrium point) *An equilibrium point of the ODE is given as*

$$\bar{y} = \left(\nu_{\bar{\rho}_1,K_1}, \pi_{\bar{\rho}_2,K_2}, \frac{\alpha_2\lambda_2}{\mu_c} \frac{1-\sum_{k=0}^{K_2} \pi_{\bar{\rho}_2,K_2}(k,0)}{\sum_{k=0}^{K_1-1} \nu_{\bar{\rho}_1,K_1}(k)} \right)$$

where $\nu_{\bar{\rho}_1,K_1}$ *is the geometric distribution on* $\{0,\dots,K_1\}$ *with parameter* $\bar{\rho}_1$, $\pi_{\bar{\rho}_2,K_2}$ *the invariant measure associated to* $Q_{\bar{\rho}_2,K_2}$ *given by* (8) *and* $(\bar{\rho}_1,\bar{\rho}_2)$ *is the solution of the fixed point equation*

$$\rho_i = \frac{1}{\lambda_i\alpha_i} \left(c_i\mu(s-E) + (q_i\mu_c - c_i\mu) \frac{\alpha_2\lambda_2}{\mu_c} \frac{1-\sum_{k=0}^{K_2} \pi_{\rho_2,K_2}(k,0)}{\sum_{k=0}^{K_1-1} \nu_{\rho_1,K_1}(k)} \right), \quad i \in \{1,2\} \tag{9}$$

with $E = \alpha_1 E_1 + \alpha_2 E_2$, E_1 *and* E_2 *being the means associated to* ν_{ρ_1,K_1} *and* π_{ρ_2,K_2}.

Proposition 2 reduces the question of existence and uniqueness of the equilibrium point to a fixed point equation. Indeed, proving the existence and uniqueness of the equilibrium point $\bar{y}$ amounts to finding a unique $(\bar{\rho}_1,\bar{\rho}_2)$ solution of the fixed point equation (9). The further analysis (existence and uniqueness) of this fixed point equation is beyond the scope of the paper. In this direction, a first approach would be to find a closed-form expression for the invariant measure π. An analytical method for generating function F associated to invariant measure π is suggested as an alternative to the probabilistic approach (see [18] for details). Some details are given in the following remark.

Remark 1 Let $\gamma_2 = \lambda_2\rho$. The global balance equation associated to π is

$$\begin{aligned} &\pi_{k,l}(\gamma_2 + \lambda_2(1-\mathbf{1}_{\{k=l=0\}}) + \delta k) \\ &= \mathbf{1}_{\{k+l<K_2\}}(\pi_{k,l+1}\lambda_2 + \pi_{k+1,l}\lambda_2\mathbf{1}_{\{l=0\}}) + \pi_{k+1,l-1}\delta(k+1)\mathbf{1}_{\{l>0\}} + \pi_{k-1,l}\gamma_2\mathbf{1}_{\{k>0\}}. \end{aligned} \tag{10}$$

Using the generating function

$$F(x,y) = \sum_{(k,l)\in\chi_2} \pi_{(k,l)} x^k y^l,$$

the global balance equation (10) yields to a functional equation on F. Although the capacity K_2 is assumed to be finite throughout the whole paper, we present here this functional equation for the case $K_2 = +\infty$ for the sake of simplicity

$$\begin{aligned} F(x,y)\left(\gamma_2(1-x) + \lambda_2\left(1-\frac{1}{y}\right)\right) \\ = F'_x(x,y)\delta(y-x) + \pi_{0,0}\lambda_2\left(1-\frac{1}{x}\right) + f(x)\lambda_2\left(\frac{1}{x}-\frac{1}{y}\right) \end{aligned}$$

where $f(x) = \sum_{k=0}^{K_1} \pi_{k,0} x^k$.

Such a functional equation is not similar to the classical case studied in [18] due to the derivative term F'_x and we do not solve it. Without an explicit form, we wonder whether the uniqueness problem could be directly solved. Such track is not explored. No closed-form solution of the invariant measure π is derived. Instead of this, the paper gives in Section 6.2 a numerical solution to equation (2).

6 Performance

In order to evaluate the impact of the incentive algorithm on the system behavior, a usual performance metric is used, i.e. the proportion of stations with no vehicle (car or gift) or no parking space available, called *problematic stations*. It characterizes how far the system is unbalanced.

Definition 1 (Performance Metric) Let $\bar{y}$ be the equilibrium point of the mean-field ODE obtained by Proposition 1. The performance metric is the limiting stationary proportion Pb of problematic stations given by

$$Pb = \alpha_1(\bar{y}_{1,0} + \bar{y}_{1,K_1}) + \alpha_2 \left(\bar{y}_{2,0,0} + \sum_{k=0}^{K_2} \bar{y}_{2,k,K_2-k} \right)$$

where K_i is the station capacity and α_i the limiting proportion of stations for cluster i, $i \in \{1, 2\}$.

The first sum in brackets is the proportion of empty and saturated stations in clusters 1, the first term $\bar{y}_{1,0}$ of stations with no car, the second term $\bar{y}_{1,K_1}$ of saturated stations in the high-demand zone. The second sum in brackets is the proportion of empty and saturated stations in cluster 2, $\bar{y}_{2,0,0}$ of stations with neither cars nor gifts and $\sum_{k=0}^{K_2} \bar{y}_{2,k,K_2-k}$ of saturated stations in the normal zone.

Optimizing the proportion of problematic stations means maximizing the number of transactions and the number of satisfied users. Our aim is to compare the performance with gifts and without gifts. The idea is to vary the fleet size parameter s, which is the limiting ratio of the total number of cars M by the total number of stations N, in order to analyze how much flexibility the gift policy gives to an operator who wants to increase the fleet size without harming the system.

6.1 Analysis of the model without gifts

From Section 5.1, the proportion of problematic stations Pb in this case is given by

$$Pb = \sum_{i=1}^{2} \alpha_i \frac{1 - \rho r_i}{1 - (\rho r_i)^{K_i+1}} (1 + (\rho r_i)^{K_i+1})$$

where $\alpha_i = \lim_{N\to\infty} N_i/N$. For $i = 1, 2$, the proportion of problematic stations in cluster i as a function of s is given by the parametric curve

$$\rho \mapsto \left(\rho\Lambda + \sum_{i=1}^{2} \alpha_i m(\nu_{\rho r_i, K_i}), \frac{1 - \rho r_i}{1 - (\rho r_i)^{K_i+1}} (1 + (\rho r_i)^{K_i+1}) \right)$$

where the first term $(1-\rho r_i)/(1-(\rho r_i)^{K_i+1})$ is the proportion of empty stations in cluster i and the second term $(\rho r_i)^{K_i+1}(1-\rho r_i)/(1-(\rho r_i)^{K_i+1})$ is the proportion of saturated stations in cluster i. As explained in Section 5.2 of [1], the proportion of problematic stations in cluster i has a minimum $2/(K_i+1)$ for ρr_i equal to 1 i.e. for $\rho = 1/r_i$. Thus, plugging in equation (4), this minimum corresponds to

$$s_i^* = \frac{\Lambda}{r_i} + \sum_{i'=1}^{2} \alpha_{i'} m(\nu_{r_{i'}/r_i, K_{i'}}).$$

where $m(\nu_{\rho,K})$ is defined by equation (5). The following result is the translation, with the notations of the paper, of the result of Section 5.2 of [1]. It gives the fleet size which corresponds to the optimal performance for a given cluster in the system without gift policy.

Proposition 3 (Optimal performance per cluster without gift policy) *For the model without gifts, the limiting stationary proportion of problematic stations in cluster $i \in \{1,2\}$ is minimal and equal to $2/(K_i+1)$ when*

$$s = s_i^* = \alpha_i \left(\frac{K_i}{2} + \frac{\lambda_i}{\mu q_i} \right) + \alpha_{3-i} \left(\frac{\gamma_{3-i}}{1-\gamma_{3-i}} - \frac{(K_{3-i}+1)\gamma_{3-i}^{K_{3-i}+1}}{1-\gamma_{3-i}^{K_{3-i}+1}} \right)$$

where $\gamma_{3-i} = (q_i \lambda_i \alpha_i)/(q_{3-i}\lambda_{3-i}\alpha_{3-i})$. The last term in brackets must be replaced by $K_{3-i}/2$ for $\gamma_{3-i} = 1$.

Note that, for $s = s_i^*$ which minimizes the proportion of problematic stations in cluster i, the proportion of problematic stations in cluster $i' \neq i$ is not optimal and is exactly $\nu_{r_{i'}/r_i, K_{i'}}(0) + \nu_{r_{i'}/r_i, K_{i'}}(K_{i'})$. Thus minimizing the problematic stations in both clusters simultaneously is not possible.

For the values of Figure 7 and $\alpha_1 = \alpha_2 = 0.5$, Proposition 3 gives $s_1^* = 29.9$ and $s_2^* = 13.1$, and for $\alpha_1 = 0.28$ and $\alpha_2 = 0.72$, $s_1^* = 21.9$ and $s_2^* = 20.4$, which can be checked in Figure 7.

Note the U-shape of the curves plotted in Figure 7. This shape is typical of these performance curves (cf [19]). Indeed, for small values of the mean number of cars per station, the proportion of empty stations is large and close to 1. Similarly, if the mean number of cars per station is large, the proportion of saturated stations is large and close to 1. Since the performance criterion includes both cases, the U-shape is observed. The contribution of empty and saturated stations to the proportion of problematic stations is illustrated by [1, Figure 2] where the proportions of empty, saturated and problematic stations are plotted.

6.2 Numerical solution

First of all, we numerically obtain the equilibrium point $\bar{y}$ of the mean-field ODE established in Proposition 1, the solution of the fixed point equation (3),

as a function of the fleet size parameter s. There are many tools to solve such an equation. We use the Anderson method implemented in Scipy, a Python library.

Figure 6 plots the performance Pb numerically obtained as a function of the fleet size parameter s, for the two-cluster model with and without gifts for a naive case: both clusters have the same number of stations, so that $\alpha_1 = \alpha_2 = 0.5$, and everyone follows the gift policy. That means the probability p that a user picks up a gift if gifts and cars are available and the probability q_1 a gift is returned to cluster 1 are equal to 1. All other parameters are given in the caption of Figure 6. We see that, for cluster 2, the cases with and without gifts are similar. But, for cluster 1, for this set of parameter values, it seems that an efficient gift policy ($p = q_1 = 1$) would allow an operator to increase the fleet size without harming the system performance and even with improving it. Indeed, for a whole range of values of the fleet size parameter s, typically $s \leq 20$, the high demand zone suffers from a lack of available cars. About 60% of the stations in the high demand zone are empty for a fleet size parameter s between 10 and 20. The effect of the incentive policy is significant in this case, since the proportion of empty stations in cluster 1 falls under 40% and even reaches 20% for $s = 20$.

Note that the crosses are simulations of the system with $N_1 = 50$ and $N_2 = 50$, the other parameters are given in Figure 6. Compared to the performance curves obtained numerically, it validates that the mean-field limit provides a good approximation for N_1 and N_2 large enough.

Figure 7 plots the performance numerically obtained for the two-cluster model with and without gifts for a more realistic case. The number of stations in the high demand zone is significantly smaller than in the normal one, the ratios are respectively $\alpha_1 = 0.28$ and $\alpha_2 = 0.72$. Figure 7 shows that the performance curves fit for small and large parameter fleet size s for both cases, with and without gifts. In between, there is a plateau where the proportion of problematic stations is close to its minimum. This implies that varying fleet size parameter s around its optimum does not degrade too much the performance which remains close to its optimum. This stability is important for the operator. The minimum proportion of problematic stations should depend on capacities K_1 and K_2, user arrival and trip rates. It is remarkable that the two plateaux correspond to the same values of s. Thus, the stations in cluster 1 do not saturate for s smaller than 30. Despite their small capacity, the high demand in cluster 1 limits the saturation.

In addition, Figure 7 shows that, for a small s, the gift policy slightly improves the performance. It is true until the two curves intersect at $s \simeq 12$. Above this value, on the plateau of cluster 1, the performance is slightly worse with the gift policy. Indeed, gifts seem to saturate cluster 1 and this slightly decreases the system performance. The mean-field approximation is again validated by simulation and the performance obtained by simulation is indicated by crosses in Figure 7.

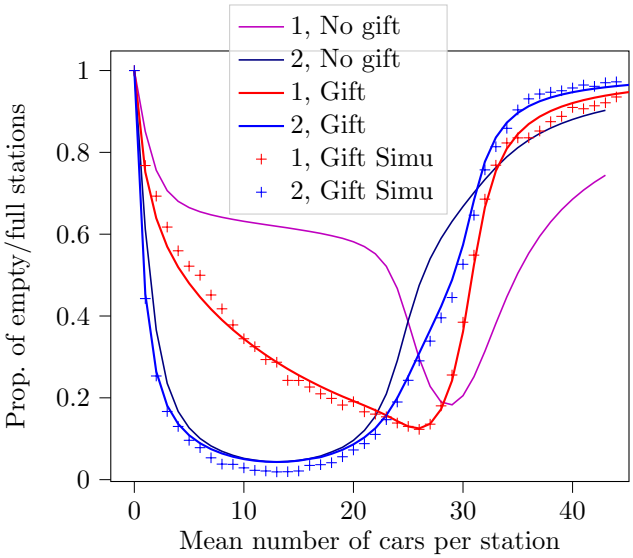

Fig. 6: Performance for both clusters (1 for the high-demand zone, and 2 for the normal zone) is numerically computed from equilibrium point equation as a function of the fleet size per station in a system with and without gifts, compared with the simulation curve in crosses, with the same number stations per cluster is $N_1 = N_2 = 50$. $K_1 = 15$, $K_2 = 45$, $\lambda_1 = 2.6$, $\lambda_2 = 1$, $\mu = \mu_c = 0.65$, $\delta = 1/14$, $c_1 = 0.5$ and $p = q_1 = 1$.

7 Discussion

7.1 Discussion of the model

Discrete Markovian framework. The exponential distributions are assumed to obtain a Markov discrete state process, i.e. the number of gifts and cars in the different stations. It is not true in real systems. This seems to be true for the arrival times of users at a station, but not for the trip times which seem heavy-tailed due to some very long trips. The behavior of the system can be affected by a log-normal trip time distribution compared to an exponential one. As to the threshold, it is deterministic in the real system. Intuitively, the exponential distribution with the same mean for the threshold should not change the behavior of the network. Large stochastic networks with general service time distributions are still largely unexplored. See Section 8 for possible approachs. The paper stays in a convenient framework.

Station-based state process. The model does not take into account the detailed moves of the cars, i.e. the fact that a car goes from station i in cluster 1 for example to station j in cluster 2. For the state of stations, these detailed moves do not matter. Mathematically, the detailed routing matrix between stations only affects the stationary behavior of the network by its invariant

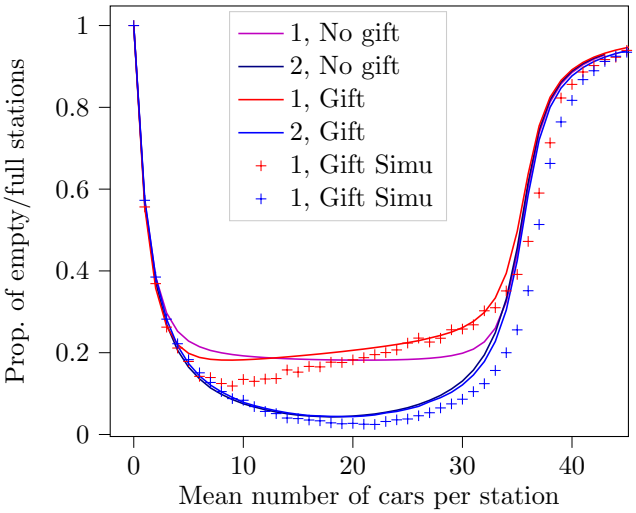

Fig. 7: Performance for both clusters (1 for the high-demand zone, and 2 for the normal zone) is numerically computed from equilibrium point equation as a function of the fleet size per station in a system with and without gifts, compared with the simulation curve in crosses, when the number of stations per cluster is given by $N_1 = 28, N_2 = 72$. $K_1 = 15$, $K_2 = 45$, $\lambda_1 = 2.6$, $\lambda_2 = 1$, $\mu = \mu_c = 0.65$, $\delta = 1/14$, $c_1 = 0.5$ and $p = q_1 = 1$.

measure. In other words, this means that the driving cars are indistinguishable and, after their departure, the origin of the trip is no longer important. Only the popularity of stations is significant, expressed as the probability that a car is dropped off at that station.

Space-homogeneity. In order to simplify the presentation, we assume that parameters do not depend on the stations. This mean-field approach can be extended to a completely heterogeneous model. It is out of the scope of the paper. See Section 8 for details.

This modeling was preceded by an analysis of real data which highlighted the low proportion of gifts offered under this incentive policy. Thus we opt for a constant probability p to choose a gift if one is available in the station. Nevertheless, a relevant option is to choose with a probability which takes into account the number of available gifts relative to cars in the station. The study is similar in this case.

Time-homogeneity. In real systems, some parameters, like the arrival rate of users, depend on time. The mathematical model does not take this into account but simulations of the time-inhomogeneous model are performed in Section 7.2.

Reservation. In real car-sharing systems, cars can be booked. It seems that such a study can still be conducted. See Section 8.

7.2 Simulations for time-inhomogeneous arrival process and real trip time distribution

We investigate now the influence of time-inhomogeneity of the parameters, especially the arrival rates of users and also the non exponential trip time distribution. Simulations are performed with arrivals according to a Poisson process with rates λ_1 and λ_2 in the stations of clusters 1 and 2, depending on time, and trip time having the distribution provided by an analysis of real data. To validate the accuracy of this time-inhomogeneous arrivals and a more realistic trip time distribution (see [13]) with Montreal FFCS system dataset, we plot in Figure 8 the average daily rate of arrivals obtained by simulation. It can be compared to that provided by data, plotted in Figure 9.

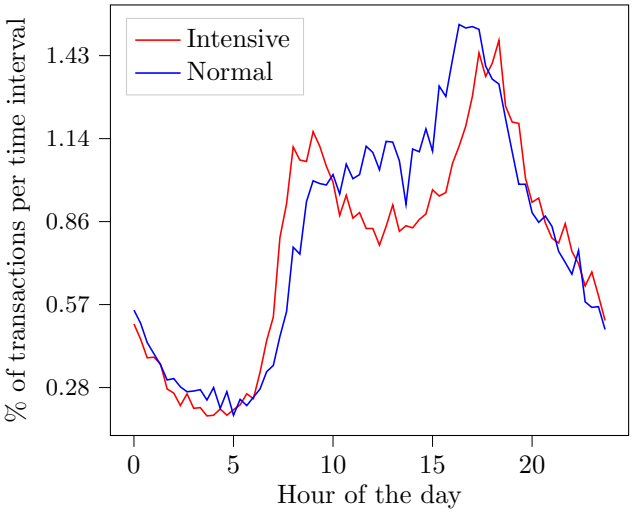

Fig. 8: Evolution of a daily arrival rate for the time-inhomogeneous simulation. Time intervals are 20 mn. The percentage is calculated based on the average total number of transactions.

Figure 10 plots the performance in both cases: time-inhomogeneous arrival rates and so-called real trip time distribution versus time-homogeneous arrival rate and trip time with exponential distribution. Performance is different for both clusters for the whole range of fleet sizes. It can be explained as follows. For large fleet sizes, more arrivals during the day and some longer trips contribute to prevent saturation in stations of both clusters in the time-inhomogeneous case. For small fleet sizes, more arrivals and larger trip times empty the stations, especially in cluster 1. It explains why the time-inhomogeneous case performs better for a mean number of cars per station around 30, but worse when it is smaller than this value.

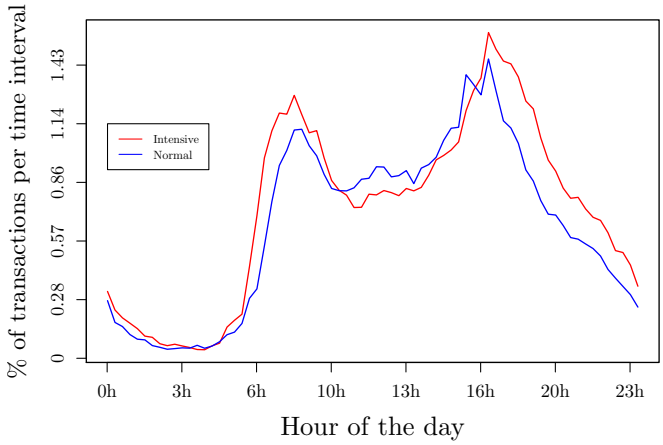

Fig. 9: Number of transactions starting in both zones during the day considering intervals of 20 mn. The percentage is calculated based on the average total number of transactions.

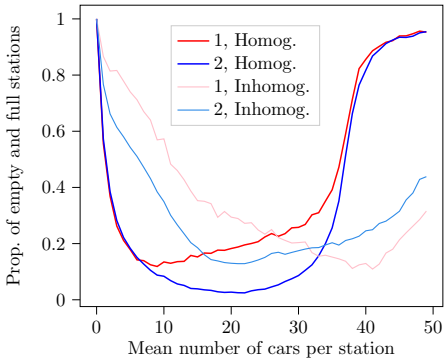

Fig. 10: Performance from a time-inhomogeneous and time-homogeneous simulation as a function of fleet size. The thin curve is plotted for time-inhomogeneous arrival rates and real trip time distribution of [13]. It is compared to the thick curve plotted for time-homogeneous arrival rates and trip times with exponential distribution of Figure 7.

In conclusion, although the homogeneous model helps us to study the influence of parameters such as δ and q, it approximates poorly time-inhomogeneous arrival rates combined with heavy-tailed trip time distribution.

8 Conclusion and future work

The analytical approach The analysis highlights an interesting random walk in the quarter-plane. Its study is important to obtain further analytical results. This issue seems to emerge in mean-field studies (see for example [20]), and involves difficult problems to prove the unique state of the stationary large-scale behavior.

Possible extensions. In Section 7.1, three extensions are highlighted.

General trip time distribution. To overcome the assumption of an exponential distribution for the trip time duration, the approachs depend on whether the trip time is bounded or not. In the first case, the approximation by an Erlang distribution seems to be natural and it allows to remain in a Markovian framework. In the second case, the main issue will be to deal with measure-valued processes.

Space-inhomogeneity. In this paper, the parameters do not depend on the stations within a zone. A possible extension is to consider different types of stations within a zone by introducing clusters. These clusters are built by grouping within a zone the stations with similar parameters. Obtaining a mean-field limit as an ODE and its numerical solution remain appropriate. It is a convenient framework for dealing with the spatial heterogeneity of real systems.

Reservation. The impact of the reservation (when a user books the car before starting the trip) on the incentive policy should be investigated. It seems technically possible with the same method. Due to the lack of reservation data in the dataset, the issue is left for future work.

The analytical approach for a related model. In our model, the dynamics of the state process leads to a *typical* station, in the sense of statistic physics, which is hard to study analytically. An alternative is to simplify the dynamics by immediatly picking up the gift for a trip when it appears. This may be relevant since the time it takes to pick it up seems negligible compared to the time it takes to become a gift. This can be checked on the data. Such a simplified version of the migration-contagion model is studied in [20] as the DOCS model where DOCS means *Departure On Change of State*, and gives analytical expressions for the invariant measure which allows to address the problem of the uniqueness of the stationary large-scale behavior. It is a work in progress.

Other performances. Another simple model in which the stations are not taken into account, but only the two zones, seems to be a coarse approximation but it is analytically tractable to obtain the ratio of gifts in the system, in order to investigate the price that the operator should pay to implement such a policy. It is a work in progress.

References

[1] Fricker, C., Gast, N., Mohamed, H.: Mean field analysis for inhomogeneous

bike sharing systems. Discrete Mathematics & Theoretical Computer Science Proceedings, 23rd Intern. Meeting on Probabilistic, Combinatorial, and Asymptotic Methods for the Analysis of Algorithms (AofA'12) (2012)

[2] Gast, N.: Expected values estimated via mean-field approximation are 1/n-accurate. Proceedings of the ACM on Measurement and Analysis of Computing Systems **1**(1), 1–26 (2017)

[3] Gast, N., Van Houdt, B.: A refined mean field approximation. Proceedings of the ACM on Measurement and Analysis of Computing Systems **1**(2), 1–28 (2017)

[4] Gast, N., Bruno, G.: A mean field model of work stealing in large-scale systems. ACM SIGMETRICS Performance Evaluation Review **38**(1), 13–24 (2010)

[5] Mukhopadhyay, A., Mazumdar, R.R., Roy, R.: Binary opinion dynamics with biased agents and agents with different degrees of stubbornness. In: 2016 28th International Teletraffic Congress (ITC 28), vol. 1, pp. 261–269 (2016). IEEE

[6] Gordon, W.J., Newell, G.F.: Closed queueing systems with exponential servers. Operations Research **15**(2), 254–265 (1967)

[7] Jackson, J.R.: Jobshop-like queueing systems. Management Science **10**(1), 131–142 (1963)

[8] Economou, A., Fakinos, D.: Product form stationary distributions for queueing networks with blocking and rerouting. Queueing Systems Theory Appl. **30**, 251–260 (1998)

[9] Fricker, C., Tibi, D.: Equivalence of ensembles for large vehicle-sharing models. The Annals of Applied Probability **27**(2), 883–916 (2017)

[10] Fayolle, G., Lasgouttes, J.-M.: Asymptotics and scalings for large product-form networks via the central limit theorem. Markov Process. Related Fields **2**, 317–348 (1996)

[11] George, D.K., Xia, C.H.: Asymptotic analysis of closed queueing networks and its implications to achievable service levels. SIGMETRICS Performance Evaluation Review **38**, 3–5 (2010)

[12] Wielinski, G., Trépanier, M., Morency, C.: Exploring service usage and activity space evolution in a free-floating carsharing service. Transportation Research Record **2673**(8), 36–49 (2019)

[13] Fricker, C., Mohamed, H., Popescu, T., Trépanier, M.: Stochastic modelling of free-floating car-sharing systems. CIRRELT (2021)

[14] Powers, D.M.W.: Applications and explanations of Zipf's law. In: New Methods in Language Processing and Computational Natural Language Learning (1998). https://aclanthology.org/W98-1218

[15] Ethier, S.N., Kurtz, T.G.: Markov Processes : Characterization and Convergence. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, New York (1986)

[16] Kelly, F.P.: Loss networks. Annals of Applied Probability **1**(3), 319–378 (1991)

[17] Robert, P.: Stochastic Networks and Queues. Applications of Mathematics: Stochastic Modelling and Applied Probability. Springer, Berlin Heidelberg (2003)

[18] Fayolle, G., Iasnogorodski, R., Malyshev, V.: Random Walks in the Quarter Plane: Algebraic Methods, Boundary Value Problems, Applications to Queueing Systems and Analytic Combinatorics. Springer, Berlin Heidelberg New-York (2003)

[19] Fricker, C., Gast, N.: Incentives and redistribution in homogeneous bike-sharing systems with stations of finite capacity. Euro journal on transportation and logistics **5**(3), 261–291 (2016)

[20] Baccelli, F., Foss, S., Shneer, V.: Migration-contagion processes. arXiv preprint arXiv:2208.03512 (2022)