The Impact of Smartphone Apps on Bike Sharing Systems

Shuang Tao

School of Operations Research and Information Engineering

Cornell University

293 Rhodes Hall, Ithaca, NY 14853

st754@cornell.edu

Jamol Pender

School of Operations Research and Information Engineering

Cornell University

228 Rhodes Hall, Ithaca, NY 14853

jjp274@cornell.edu

April 19, 2020

Abstract

Bike-sharing systems are exploding in cities around the world as more people are adopting sustainable transportation solutions for their everyday commutes. However, as more people use the system, riders often encounter that bikes or docks might not be available when they arrive to a station. As a result, many systems like CitiBike and Divvy provide riders with information about the network via smartphone apps so that riders can find stations with available bikes. However, not all customers have adopted the use of these smartphone apps for their station selection. By combining customer choice modeling and finite capacity queueing models, we explore the impact of the smartphone app technology to increase throughput and reduce blocking in bike sharing systems. To this end, we prove a mean-field limit and a central limit theorem for an empirical process of the number of stations with k bikes. We also prove limit theorems for a new process called the ratio process, which characterizes the proportion of stations whose bike availability ratio lies within a particular partition of the interval [0,1]. For the mean field limit, we prove that the equilibrium exists, is unique, and that the stationary distribution of the empirical measure converges to a Dirac mass at the same equilibrium, thus showing an interchange of limits result ($\lim_{t\to\infty} \lim_{N\to\infty} = \lim_{N\to\infty} \lim_{t\to\infty}$). Our limit theorems provide insight on the mean, variance, and sample path dynamics of large scale bike-sharing systems. Our results illustrate that if we increase the proportion of customers that use smartphone app information, the entropy of the bike sharing network is reduced, and riders experience less blocking in the network.

1 Introduction

Bike sharing is expanding rapidly around the world. Around 1000 cities worldwide have implemented bike sharing programs and the majority of them are achieving great success. Bike sharing programs aim to provide affordable short-term trips for people living in urban areas as an alternative to other public transportation or private vehicles. They have the benefits of reducing traffic congestion, noise, and mitigating air pollution. Studies also report observing an increase in cycling and health benefits where bike sharing systems are run.

Bike sharing has developed significantly since its initial launch in Amsterdam in 1965, and it is generally considered that bike sharing has gone through three generations of major changes in the implementation and design of the system. The contemporary 3rd generation schemes, which are also called information technology (IT) based systems, exploited information and communication technology to effectively automate the bike sharing system. This includes automated smartcards (or magnetic stripe card), electronic bike locking, electronic payment, GPS (Global Positioning System) tracking of bikes, and networked self-service stations managed by central computer systems and technologies like Radio Frequency Identification (RFID) to monitor the location of bikes in the system. Many systems also introduced the use of websites or smartphone apps to provide users with real-time information about available bikes and open docks at nearby stations.

The biggest issue for all bike sharing systems is the scarcity of resources to move all riders around each city at all times of the day. That is, customers can encounter empty stations when picking up bikes, and they can also encounter full stations when dropping off bikes. In both situations, customers get blocked and therefore system throughput is reduced and customers are frustrated and might leave the system altogether. Some empirical and analytical studies of bike availability at CitiBike stations show that during peak hours, over 30% of the stations are empty and 8% of the stations are full Faghih-Imani et al. [8], Freund et al. [10], Ghosh et al. [13], Hampshire and Marla [14], Kabra et al. [16], Jian et al. [15], Tao and Pender [30]. The pressing need to come up with economic yet effective solutions to this issue is there.

The most direct way is real-time rebalancing of bikes across stations. Examples of papers that analyze real-time rebalancing strategies are Erdoğan et al. [6], Contardo et al. [4], Freund et al. [10], Ghosh et al. [13], Nair and Miller-Hooks [18], O'Mahony [23], O'Mahony and Shmoys [22], Schuijbroek et al. [28], Raviv et al. [27], Faghih-Imani et al. [9], Singla et al. [29], Freund et al. [11]. However, in practice, real-time rebalancing is both expensive and time-consuming, and it has not been adapted by most large scale bike sharing systems. The main reason is that trucks need to pick up bikes from nearly full stations and deliver them to empty stations during high demand times. Since peak hours in a workday only last about 1-2 hours, real-time rebalancing may not be able to solve unbalanced inventory in bike sharing systems. Another solution is to implement real-time dynamic pricing, see for example Fricker and Gast [12], Kabra et al. [16], Pfrommer et al. [26], Waserhole et al. [32], Waserhole and Jost [31], Chemla et al. [2]. However, given the majority of the users of CitiBike are subscribers who pay an annual fee to get unlimited rides, real-time pricing will not make a big difference since it only affects non-subscribers.

With the advances in technology, many large scale bike sharing systems such as CitiBike

now give their users the ability to see real-time bike/dock availability at nearby stations through a smartphone app and web API. A picture of the CitiBike app is given in Figure 1. This information helps users make better decisions such as where to pick up bikes from nearby locations. This means that when stations are empty or have very few bikes, users who have this information through smartphone app will be less likely to pick up bikes at these stations. Similarly, when stations are full or nearly full, users who have the information are more likely to pick up bikes at these stations. Intuitively, providing users real-time information about bike/dock availability should have the power of guiding users to move bikes from stations with more bikes to stations with less bikes and therefore making the system more balanced. However, it is not known how the app itself affects the underlying network dynamics since this additional information can change the behavior of the users.
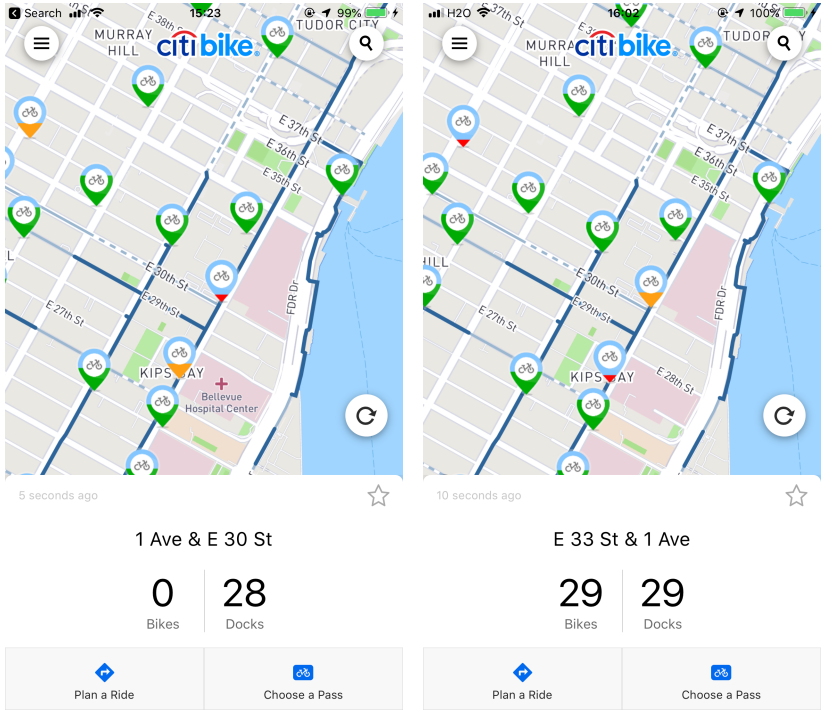

Figure 1: CitiBike smartphone app screenshots

In this paper, we study the influence of smartphone app and its power to guide users to non-empty stations. Inspired by the empirical work of Kabra et al. [16], Nirenberg et al. [19] and the theoretical work of Pender et al. [24, 25], Novitzky et al. [21, 20], our approach combines customer choice models and finite capacity queueing models. However, our work is more analytical in our approach since we prove a mean-field and central limit theorems for the empirical process of the proportion of stations with k bikes. These limit theorems allow us to study the impact of information by analyzing the asymptotic behavior of the mean, variance and sample path dynamics of large scale bike sharing systems under different choice models and different parameters, such as the proportion of customers that use the smartphone app information, customers' sensitivity to information, etc. Moreover, we also develop a mean-field and central limit theorems for the bike availability ratio process, which counts the proportion of stations whose bike availability falls into an interval $[i/n, (i + 1)/n)$. This

ratio process is especially useful when station capacities are not uniform and identical. For example, in the case when we have two stations with capacity equal to 5 and 70 respectively. Having 3 bikes at station 1 means it already achieves over half of the capacity, while at station 2 it clearly means too few bikes. In such case the ratio process will provide more accurate information about the condition of a station than the empirical process. Finally, for both the empirical process and the ratio process, we study the system in steady state and show that an equilibrium exists and is unique. We also show that the steady state empirical process and ratio process converge to a Dirac mass at the equilibrium point, thus proving an interchanging of limits result.

1.1   Main Contributions of Paper

The contributions of this work can be summarized as follows:

- We prove a transient mean field and central limit theorem for the empirical process.

- We construct a new process called the ratio process and derive a mean field and central limit theorem for the ratio process and compute its steady state equilibrium.

- We prove an interchange of limits for the mean field limit empirical process and the ratio process. We show that the limit is a Dirac mass.

- We show that our stochastic choice queueing model can replicate empirical process and ratio process given by the data of the CitiBike network.

1.2   Organization of Paper

The remainder of this paper is organized as follows. Section 2 introduces our bike sharing model with customer choice. We also prove a mean field and central limit theorem for the empirical process for the proportion of stations with k bikes. In Section 3, we also prove a mean field and central limit theorem for the empirical process for the proportion of stations whose bike availability falls into a interval $[i/n, (i+1)/n)$. In Section 4, we compute explicitly the equilibrium and Lyapunov functions of the mean field limit of the empirical process and the ratio process. Finally, we show that both the empirical process and the ratio process satisfy an interchange of limits. In Section 5, we discuss the simulation results of our model, in both stationary and non-stationary arrival processes settings. These results incorporates real-life data from CitiBike, and provide insights into how different aspects of the impact of information affect the behavior of empirical measure over time. In Section 6, we give concluding remarks and provide some future directions of research. Finally, in Section 7 the Appendix, additional proofs and theorems are given.

2   Stochastic Model and Limit Theorems

In this section, we construct a Markovian bike-sharing queueing model where customers can pick-up and drop-off bikes at each station if there is available capacity.

2.1 Bike Sharing Model with Customer Choice Dynamics

Motivated by the CitiBike system, we consider a bike sharing system with N stations and a fleet of M bikes in total. We assume that the arrival of customers to the stations are independent Poisson processes with rate λ_i for the ith station. When a customer arrives at a station, if there is no bike available, he will then leave the system and is immediately blocked and lost. Otherwise, he takes a bike and rides to station j with probability P_j. We assume that the travel time of the rider is exponentially distributed with mean $1/\mu$, for every transition from one station to another. Since we are concerned with finite capacity stations, we assume that the ith station has a bike capacity of K_i, which is assumed to be finite for all stations. Thus, when a customer arrives at station j, if there are less than K_j bikes in this station, he returns his bike and leaves the system. If there are exactly K_j bikes (i.e. the station is full), the customer randomly chooses another station k with probability P_k and goes to that station to drop the bike off. As before, it takes a time that is exponentially distributed with mean $1/\mu$. Finally, the customer rides like this again until he can return his bike to a station that is not full.

Below we provide Table 1 for the reader's convenience to so that they understand the notation that we will use throughout the rest of the paper.

Table 1: Notation

N	Number of stations
M	Number of bikes
K_i	Capacity at station i
λ_i	Arrival rate at station i
$1/\mu$	Mean travel time
P_i	Routing probability to station i
$X_i(t)$	Number of bikes at station i at time t
γ	Average number of bikes at each station
p	The fraction of users that have information about bike distribution
g	The choice function
θ	Users' sensitivity to information (exponential choice function)
c	Users' sensitivity threshold to information (minimum choice function)
α	Users' sensitivity to information (polynomial choice function)

Throughout the rest of the paper, we assume without loss of generality that the arrival rate $\lambda_i = \lambda$ and service rate $\mu = 1$, and that the routing probability from each station is uniform i.e. $P_i = 1/N$. We also assume that capacities across all stations are equal, i.e. $K_i = K, i = 1, ..., N$. We refer to Tao and Pender [30] on how our model can be extended seamlessly to settings with non-uniform arrival rate, routing probabilities and capacities. With our notation in hand, we are ready to develop our stochastic model for our bike-sharing network.

Let $\mathbf{X}(t) = (X_1(t), \cdots, X_N(t))$, where $X_i(t)$ is the number of bikes at station i at time t, then $X_i(t)$ is a continuous time Markov chain (CTMC), in particular a state dependent $M/M/1/K_i$ queue. In this model, the rate of dropping off bikes at station i is equal to

$\mu P_i \left(M - \sum_{k=1}^{N} X_k(t) \right) \mathbf{1}\{X_i(t) < K_i\}$, and the rate of retrieving bikes at station i is equal to $\left((1-p)\lambda + p\lambda_i N \frac{g(X_i(t))}{\sum_{j=1}^{N} g(X_j(t))} \right) \mathbf{1}\{X_i(t) > 0\}$, where g is an arbitrary non-decreasing function that is also non-negative.

Proposition 2.1. *For any continuous function f, the CTMCs $\{X_i(t)\}_{i=1}^{N}$ satisfy the following functional forward equations,*

$$
\overset{\bullet}{\mathbb{E}}[f(X_i(t))|X_i(0) = x_i] \equiv \frac{d}{dt}\mathbb{E}[f(X_i(t))|X_i(0) = x_i]
$$

$$
= \mathbb{E}\left[(f(X_i(t)+1) - f(X_i(t))) \cdot \left(\mu P_i \left(M - \sum_{k=1}^{N} X_k(t) \right) \mathbf{1}\{X_i(t) < K_i\} \right) \right]
$$

$$
+ \mathbb{E}\left[(f(X_i(t)-1) - f(X_i(t))) \cdot \left(\left((1-p)\lambda + p\lambda N \frac{g(X_i(t))}{\sum_{j=1}^{N} g(X_j(t))} \right) \mathbf{1}\{X_i(t) > 0\} \right) \right].
$$

$$(2.1)$$

2.2 Empirical Bike Process

Although the functional forward equations given in Proposition 2.1 describe the exact dynamics of the mean of the bike sharing system, the system of differential equations are not closed. This non-closure property of the functional forward equations in this model arises from the fact that the bike sharing system has finite capacity. More importantly, it also implies that we need to know a priori the full distribution of the whole stochastic process $\mathbf{X}(t)$ in order to calculate the mean or variance or any moment for that matter. Finally if we even wanted to solve these equations and knew the entire distribution of $\mathbf{X}(t)$ a priori, there are still $O(N^2)$ differential equations that would need to be numerically integrated. For some major bike sharing networks such as CitiBike in NYC, N is around 900; for Divvy in Chicago, N is around 600. This can be very computationally expensive and thus, we must take a different approach to analyze our bike sharing system.

Moreover, if we want to analyze the limiting behavior of $\{X_i\}_{i=1}^{N}$ as CTMCs, as we let N go to ∞, the mean field limit would become infinite dimensional, which is quite complicated. However, if we instead analyze an empirical measure process for $\mathbf{X}(t)$, we can use the finite capacity nature of the bike sharing system to our advantage and have a finite dimensional CTMC for the empirical measure process. See Tao and Pender [30] for more discussion on this topic.

Thus, we construct an empirical measure process that counts the proportions of stations with n bikes. This empirical measure process is given below by the following equation:

$$
Y_t^N(n) = \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}\{X_i(t) = n\}.
$$

$$(2.2)$$

Conditioning on $Y_t^N(n) = y(n)$, the transition rates of y are specified as follows:

- When there is an customer arrival to a station with n bikes and relative utilization r to retrieve a bike, the proportion of stations having n bikes goes down by $1/N$, the

proportion of stations having $n-1$ bikes goes up by $1/N$, and the transition rate Q^N is

$$Q^N\left(y, y + \frac{1}{N}(\mathbf{1}_{n-1} - \mathbf{1}_n)\right) = \left((1-p)\lambda N y(n) + p\lambda N \frac{y(n)g(n)}{\sum_{j=0}^K y(j)g(j)}\right)\mathbf{1}_{n>0}. \tag{2.3}$$

- When there is a customer returning a bike to a station with n bikes and relative utilization r, the proportion of stations having n bikes goes down by $1/N$, the proportion of stations having $n+1$ bikes goes up by $1/N$, and the transition rate Q^N is

$$Q^N\left(y, y + \frac{1}{N}(\mathbf{1}_{n+1} - \mathbf{1}_n)\right) = y(n)\cdot\mu\cdot\left(M - \sum_{n'} n'y(n')N\right)\mathbf{1}_{n<K}$$

$$= y(n)N\left(\frac{M}{N} - \sum_{n'} n'y(n')\right)\mathbf{1}_{n<K}. \tag{2.4}$$

Similarly, we have the functional forward equations for $Y^N(t)$ as follows.

Proposition 2.2. *For any continuous function f, $Y^N(t)$ satisfies the following functional forward equation,*

$$\overset{\bullet}{\mathbb{E}}[f(Y_n^N(t))|Y_n^N(0) = y_n(0)] \equiv \frac{d}{dt}\mathbb{E}[f(Y_n^N(t))|Y_n^N(0) = y_n(0)]$$

$$= \mathbb{E}\left[\left(f\left(Y_n^N(t) + 1/N\right) - f(Y_n^N(t))\right)\cdot\left(Y_{n+1}^N(t)N\lambda\left((1-p) + p\frac{g(n+1)}{\sum_{j=0}^K Y_j^N(t)g(j)}\right)\mathbf{1}_{n<K}\right.\right.$$

$$\left.\left. + Y_{n-1}^N(t)N\left(\frac{M}{N} - \sum_{n'} n'Y_{n'}^N(t)\right)\mathbf{1}_{n>0}\right)\right]$$

$$+ \mathbb{E}\left[\left(f(Y_n^N(t) - 1/N) - f(Y_n^N(t))\right)\cdot\left(Y_n^N(t)N\lambda\left((1-p) + p\frac{g(n)}{\sum_{j=0}^K Y_j^N(t)g(j)}\right)\mathbf{1}_{n>0}\right.\right.$$

$$\left.\left. + Y_n^N(t)N\left(\frac{M}{N} - \sum_{n'} n'Y_{n'}^N(t)\right)\mathbf{1}_{n<K}\right)\right], \tag{2.5}$$

for $n = 0, 1, ..., K$.

However, we should mention that the functional forward equations in Proposition 2.2 are also not closed. Therefore, we need to use another method for the analysis of this model to get around the non-closure properties of the equations, which leads us to the mean field limit and diffusion limit analysis in the following sections.

2.3 Mean Field Limit of Empirical Process

In this section, we prove the mean field limit for our bike sharing model. A mean field limit describes the large station dynamics of the bike sharing network over time. Deriving the

mean field limit allows us to gain insights on average system behaviors, when the the demand for bikes and the number of stations are very large. Thus, we avoid the need to study an N-dimensional CTMC and compute its steady state distribution in this high dimensional setting.

Theorem 2.3 (Functional Law of Large Numbers). *Let $|.|$ denote the Euclidean norm in $\mathbb{R}^{K+1}$. Suppose that $\lim_{N \to \infty} \frac{M}{N} = \gamma$ and $Y_0^N \overset{p}{\to} y_0$, then we have for $\forall \epsilon > 0$*

$$\lim_{N \to \infty} P \left(\sup_{t \le t_0} |Y_t^N - y_t| > \epsilon \right) = 0,$$

where y_t is the unique solution to the following differential equation starting at y_0

$$\overset{\bullet}{y} = b(y), \tag{2.6}$$

where $b : [0,1]^{K+1} \to \mathbb{R}^{K+1}$ is a vector field satisfies

$$
\begin{aligned}
b(y) \quad = \quad & \sum_{n=0}^{K} \Bigg[\left((1-p) + p \frac{g(n)}{\sum_{j=0}^{K} y_j g(j)} \right) (\mathbf{1}_{n-1} - \mathbf{1}_n) \lambda \mathbf{1}_{n>0} \\
& + \left(\gamma - \sum_{j=0}^{K} j y_j \right) (\mathbf{1}_{n+1} - \mathbf{1}_n) \mathbf{1}_{n<K} \Bigg] y_n,
\end{aligned}
\tag{2.7}
$$

or

$$
b(y_0) = \quad \underbrace{- \left(\gamma - \sum_{j=0}^{K} j y_j \right) y_0}_{\text{return a bike to a no-bike station}} \quad + \underbrace{\left((1-p) + p \frac{g(1)}{\sum_{j=0}^{K} y_j g(j)} \right) \lambda y_1}_{\text{retrieve a bike from a 1-bike station}},
\tag{2.8}
$$

$$
b(y_k) = \underbrace{\left((1-p) + p \frac{g(k+1)}{\sum_{j=0}^{K} y_j g(j)} \right) \lambda y_{k+1}}_{\text{retrieve a bike from a } k+1\text{-bike station}} - \underbrace{\left(\left((1-p) + p \frac{g(k)}{\sum_{j=0}^{K} y_j g(j)} \right) \lambda + \gamma - \sum_{j=0}^{K} j y_j \right) y_k}_{\text{retrieve and return a bike to a } k\text{-bike station}}
$$

$$
+ \underbrace{\left(\gamma - \sum_{j=0}^{K} j y_j \right) y_{k-1}}_{\text{return a bike to a } k-1\text{-bike station}} \quad ,
$$

$$\tag{2.9}$$

for $k = 1, ..., K-1$, and

$$
b(y_K) = \underbrace{- \left((1-p) + p \frac{g(K)}{\sum_{j=0}^{K} y_j g(j)} \right) \lambda y_K}_{\text{retrieve a bike from a } K\text{-bike station}} + \underbrace{\left(\gamma - \sum_{j=0}^{K} j y_j \right) y_{K-1}}_{\text{return a bike to a } K-1\text{-bike station}} \quad .
\tag{2.10}
$$

Proof. The complete proof is provided in the Appendix (Section 7). $\qquad\square$

Our mean field limit shows that the analysis of our bike sharing system is essentially solving a system of ordinary differential equations when the number of stations N goes to infinity. Using the fluid limit, we can estimate important performance measures such as proportion of empty or full stations, what factors (such as fleet size and capacity) are affecting these performance measures, and what we can do to improve the system performance. Deriving the mean field limit also allows us to compute the system's equilibrium point, which is addressed in Section 4.

However, knowing the mean field limit is not enough, we would like to know more about the stochastic variability of the system, i.e. the fluctuations around the mean field limit, and what is stochastic sample path behavior of the system. Thus, in the subsequent section we develop a functional central limit theorem for our bike sharing model, and explain why it is important for understanding stochastic fluctuations.

2.4 Diffusion Limit of Empirical Process

In this section, we derive the diffusion limit of our stochastic empirical process bike sharing model. Diffusion limits are critical for obtaining a deep understanding of the sample path behavior of stochastic processes. One reason is that diffusion limits describe the fluctuations around the mean or mean field limit and can help understand the variance or the asymptotic distribution of the stochastic process being analyzed, and help provide confidence intervals to the mean field limit. We define our diffusion scaled bike sharing model by subtracting the mean field limit from the scaled stochastic process and rescaling by $\sqrt{N}$. Thus, we obtain the following expression for the diffusion scaled bike sharing empirical process

$$D_t^N = \sqrt{N}(Y_t^N - y_t). \tag{2.11}$$

Unlike many other ride-sharing systems such as Lyft or Uber, bike sharing does not have a pricing mechanism to redistribute bikes in real-time to satisfy demand. For this reason, it is essential to understand the dynamics and behavior of D_t^N. D_t^N can be useful for describing the probability that the proportion of stations with i bikes exceeds a threshold i.e. $\mathbb{P}(Y_t^N > x)$ for some $x \in [0,1]^{K+1}$. It also describes this probability in a situation where there is no control or rebalancing of bikes in the system. This knowledge of the uncontrolled system is especially important for newly-started bike sharing systems who are still in the process of gathering information about the system demand. Having the diffusion limit help managers of these bike sharing systems understand the system dynamics and stability, which in terms help them make long term managerial decisions. It is also helpful in the case when the operators of the bike sharing system have no money for rebalancing the system to meet real-time demand.

Using the semi-martingale decomposition of Y_t^N given in Equation (7.55), we can write a similar decomposition for D_t^N as follows:

$$\begin{aligned}
D_t^N &= \sqrt{N}(Y_0^N - y_0) + \sqrt{N}M_t^N + \int_0^t \sqrt{N}[\beta(Y_s^N) - b(y_s)]ds \\
&= D_0^N + \sqrt{N}M_t^N + \int_0^t \sqrt{N}[\beta(Y_s^N) - b(Y_s^N)]ds + \int_0^t \sqrt{N}[b(Y_s^N) - b(y_s)]ds.
\end{aligned} \tag{2.12}$$

Define

$$D_t = D_0 + \int_0^t b'(y_s) D_s ds + M_t, \tag{2.13}$$

where $b'(y) = \left(\frac{\partial b(y(i))}{\partial y(j)}\right)_{ij}$ and $M_t = (M_t(0), \cdots, M_t(K)) \in \mathbb{R}^{K+1}$ be a real continuous centered Gaussian martingale, with Doob-Meyer brackets given by

$$\langle M(k)\rangle_t = \int_0^t (b_+(y_s)(k) + b_-(y_s)(k)) ds,$$

$$\langle M(k), M(k+1)\rangle_t = -\int_0^t \left[\left(\lambda(1-p) + p\frac{g(k+1)}{\sum_{j=0}^K y_j g(j)}\right) y_{k+1} \right.$$

$$\left. + \left(\gamma - \sum_j j y_j\right) y_k\right] ds \quad \text{for } k < K,$$

$$\langle M(k), M(j)\rangle_t = 0 \quad \text{for } |k-j| > 1. \tag{2.14}$$

Here $b_+(y), b_-(y)$ denote the positive and the negative parts of function $b(y)$ respectively.

Now we state the functional central limit theorem for the empirical measure process below.

Theorem 2.4 (Functional Central Limit Theorem). *Consider D_t^N in $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^{K+1})$ with the Skorokhod J_1 topology, and suppose that $\limsup_{N\to\infty} \sqrt{N}(\frac{M}{N} - \gamma) < \infty$. Then if D_0^N converges in distribution to D_0, then D_t^N converges to the unique OU process solving*

$$D_t = D_0 + \int_0^t b'(y_s) D_s ds + M_t$$

in distribution, where $b'(y)$ is specified as follows,

$$\frac{\partial}{\partial y_i} b(y_0) = iy_0 - \left(\gamma - \sum_{j=0}^K j y_j\right) \mathbf{1}_{i=0} + \lambda\left[(1-p) + p\frac{g(1)}{\sum_{j=0}^K y_j g(j)}\right] \mathbf{1}_{i=1} - \frac{\lambda p y_1 \cdot g(1) g(i)}{\left(\sum_{j=0}^K y_j g(j)\right)^2}, \tag{2.15}$$

$$\frac{\partial}{\partial y_i} b(y_k) = -\frac{\lambda p \cdot g(i)}{\left(\sum_{j=0}^K y_j g(j)\right)^2} (y_{k+1} g_{k+1}(k+1) - y_k g(k)) + i(y_k - y_{k-1})$$

$$+ \lambda\left(1 - p + p\frac{g_{k+1}(k+1)}{\sum_{j=0}^K y_j g(j)}\right) \mathbf{1}_{i=k+1} + \left(\gamma - \sum_{j=0}^K j y_j\right) \mathbf{1}_{i=k-1}$$

$$- \left[\left(1 - p + p\frac{g(k)}{\sum_{j=0}^K y_j g(j)}\right)\lambda + \gamma - \sum_{j=0}^K j y_j\right] \mathbf{1}_{i=k}, \quad \text{for } k = 1, ..., K-1, \tag{2.16}$$

$$\frac{\partial}{\partial y_i} b(y_K) = \frac{\lambda p y_K \cdot g(K) g(i)}{\left(\sum_{j=0}^K y_j g(j)\right)^2} - iy_{K-1} - \left(1 - p + p\frac{g(K)}{\sum_{j=0}^K y_j g(j)}\right)\lambda \mathbf{1}_{i=K} + \left(\gamma - \sum_{j=0}^K j y_j\right) \mathbf{1}_{i=K-1}. \tag{2.17}$$

Proof. To prove Theorem 2.4, we take the following 4 steps.

1). (Lemma 7.4) $\sqrt{N}M_t^N$ is a family of martingales independent of D_0^N with Doob-Meyer brackets given by

$$\left\langle \sqrt{N}M^N(k) \right\rangle_t = \int_0^t (\beta_+(Y_s^N)(k) + \beta_-(Y_s^N)(k))ds,$$

$$\left\langle \sqrt{N}M^N(k), \sqrt{N}M^N(k+1) \right\rangle_t = -\int_0^t \left[\left((1-p) + p\frac{g(k+1)}{\sum_{j=0}^K jy_j} \right) \lambda Y_s^N(k+1) \right.$$
$$\left. + \left(\frac{M}{N} - \sum_{j=0}^K jY_s^N(j) \right) Y_s^N(k) \right] ds \quad \text{for } k < K,$$

$$\left\langle \sqrt{N}M^N(k), \sqrt{N}M^N(j) \right\rangle_t = 0 \quad \text{for } |k-j| > 1. \tag{2.18}$$

2). (Lemma 7.6) For any $T \geq 0$,

$$\limsup_{N\to\infty} \mathbb{E}(|D_0^N|^2) < \infty \Rightarrow \limsup_{N\to\infty} \mathbb{E}(\sup_{0\leq t\leq T} |D_t^N|^2) < \infty.$$

3). (Lemma 7.7) If $(D_0^N)_{N=1}^\infty$ is tight then $(D^N)_{N=1}^\infty$ is tight and its limit points are continuous.

4). If D_0^N converges to D_0 in distribution, then D_t^N converges to the unique OU process solving $D_t = D_0 + \int_0^t b'(y_s)D_s ds + M_t$ in distribution.

We provide the proofs of Lemma 7.4, Lemma 7.6 and Lemma 7.7 (step 1-3) in the Appendix (Section 7). For step 4), by Theorem 4.1 in Chapter 7 of Ethier and Kurtz [7], it suffices to prove the following condition holds

$$\sup_{t\leq T} \left| \int_0^t \left\{ \sqrt{N}[b(Y_s^N) - b(y_s)] - b'(y_s)D_s^N \right\} ds \right| \xrightarrow{p} 0.$$

By Equations (2.15), (2.16), and (2.17), we know that $b(y_t)$ is continuously differentiable with respect to y_t. By the mean value theorem, for every $0 \leq s \leq t$ there exists a vector Z_s^N in between Y_s^N and y_s such that

$$b(Y_s^N) - b(y_s) = b'(Z_s^N)(Y_s^N - y_s).$$

Therefore

$$\int_0^t \left\{ \sqrt{N}[b(Y_s^N) - b(y_s)] - b'(y_s)D_s^N \right\} ds = \int_0^t [b'(Z_s^N) - b'(y_s)]D_s^N ds.$$

We know that

$$\lim_{N\to\infty} \sup_{t\leq T} |b'(Z_s^N) - b'(y_s)| = 0 \quad \text{in probability}$$

by the mean field limit convergence and the uniform continuity of b'. By applying Chebyshev's inequality we have that D_s^N is bounded in probability. Thus, by Lemma 5.6 in Ko and Pender [17],

$$\sup_{t \le T} \left| \int_0^t \left\{ \sqrt{N}[b(Y_s^N) - b(y_s)] - b'(y_s)D_s^N \right\} ds \right| \xrightarrow{p} 0.$$

$\square$

3 Ratio Process Perspective

In this section, we explore a new empirical measure process called the ratio process, which is the empirical measure of the ratio between the number of bikes at a station and its capacity, when station capacities differs among stations. According to data from CitiBike, station capacities range from 10 docks to 60 docks. In this case, the ratio process provides a better understanding of the bike availability across the network than the empirical measure process mentioned in previous section, especially for understanding full stations when station capacities differ. Let $K_{\max}$ be the maximum capacity among stations, we define the ratio process R_t^N as follows,

$$R_t^N = \left(R_t^N(0), R_t^N(1), \cdots, R_t^N(K_{\max} - 1) \right), \tag{3.19}$$

where for $j = 0, \cdots, K_{\max} - 1$,

$$R_t^N(j) = \frac{1}{N} \sum_{i=1}^{N} \mathbf{1} \left\{ \frac{j}{K_{\max}} \le \frac{X_i(t)}{K_i} < \frac{j+1}{K_{\max}} \right\}. \tag{3.20}$$

To demonstrate the difference between the empirical measure process and the ratio process in real life setting, we collected CitiBike's real-time system data in the General Bikeshare Feed Specification (GBFS) format. GBFS provides a real-time snapshot of the city's fleet, which includes the number of bikes and docks at every CitiBike station at the time. The data was pulled from CitiBike's API once every second for a period of time in September 2017. Figure 2 gives a screenshot of the empirical measure process of bike distribution and its ratio process at a given time. We can observe that when capacities differ across stations, the ratio process is indeed better at capturing the proportion of stations at full capacity.

It is easy to show that R_t^N is a Markov process with the following transition rates:

- When a customer picks up a bike at station i,

$$Q^N \left(R_t^N, R_t^N + \frac{1}{N} \left(\mathbf{1}_{\left\lfloor \frac{(X_i(t)-1)K_{\max}}{K_i} \right\rfloor} - \mathbf{1}_{\left\lfloor \frac{X_i(t)K_{\max}}{K_i} \right\rfloor} \right) \right)$$
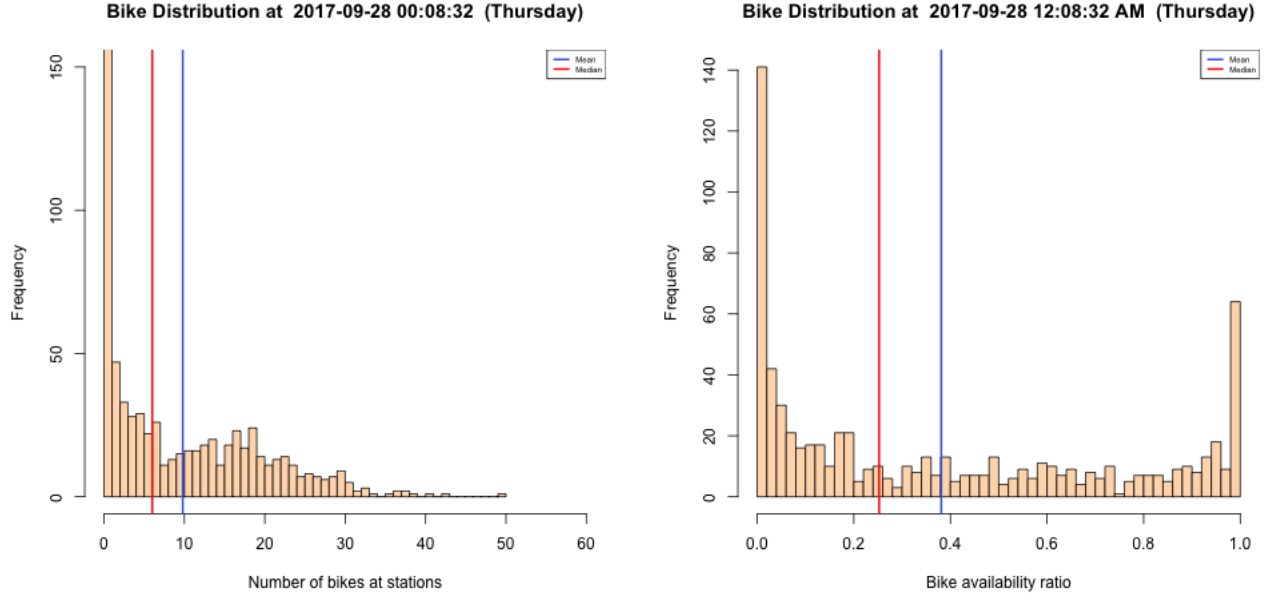$$= \left((1-p)\lambda_i + p\lambda_i N \frac{g(X_i(t))}{\sum_{j=1}^{N} g(X_j(t))} \right) \mathbf{1}\{X_i(t) > 0\}. \tag{3.21}$$

12

Figure 2: A Snapshot of CitiBike Empirical Measure Process vs. Ratio Process

- When a customer drops off a bike at station i,

$$Q^N \left(R_t^N, R_t^N + \frac{1}{N} \left(\mathbf{1}_{\left\lfloor \frac{(X_i(t)+1)K_{\max}}{K_i} \right\rfloor} - \mathbf{1}_{\left\lfloor \frac{X_i(t)K_{\max}}{K_i} \right\rfloor} \right) \right)$$

$$= \mu P_i \left(M - \sum_{k=1}^{N} X_k(t) \right) \mathbf{1}_{X_i(t) < K_i}. \tag{3.22}$$

We can prove the following functional forward equations,

$$\overset{\bullet}{\mathbb{E}}[f(R_t^N)|R_0^N = r_0^N] \equiv \frac{d}{dt} \mathbb{E}[f(R_t^N)|R_0^N = r_0^N]$$

$$= \sum_{i=1}^{N} \mathbb{E}\left[\left(f(R_t^N + \frac{1}{N} \left(\mathbf{1}_{\left\lfloor \frac{(X_i(t)+1)K_{\max}}{K_i} \right\rfloor} - \mathbf{1}_{\left\lfloor \frac{X_i(t)K_{\max}}{K_i} \right\rfloor} \right)) - f(R_t^N) \right) \cdot (\mu P_i (M$$

$$- \sum_{k=1}^{N} X_k(t) \right) \mathbf{1}\{X_i(t) < K_i\}) \Bigg]$$

$$+ \sum_{i=1}^{N} \mathbb{E}\left[\left(f(R_t^N + \frac{1}{N} \left(\mathbf{1}_{\left\lfloor \frac{(X_i(t)-1)K_{\max}}{K_i} \right\rfloor} - \mathbf{1}_{\left\lfloor \frac{X_i(t)K_{\max}}{K_i} \right\rfloor} \right)) - f(R_t^N) \right) \cdot (((1-p)\lambda$$

$$+ p\lambda N \frac{g(X_i(t))}{\sum_{j=1}^{N} g(X_j(t))} \right) \mathbf{1}\{X_i(t) > 0\}) \Bigg]. \tag{3.23}$$

The downside of this representation of R_t^N is that the state transition depends on $X_i(t)$ in a very complicated way. Thus, another way to deal with the ratio process is through the

empirical measure process $\tilde{Y}_t^N(n,k)$, which is the proportion of stations with n bikes and capacity k at time t, defined as

$$\tilde{Y}_t^N(n,k) = \frac{1}{N}\sum_{i=1}^{N}\mathbf{1}\{X_i(t) = n, K_i = k\}. \tag{3.24}$$

Condition on $\tilde{Y}_t^N(n,k) = \tilde{y}(n,k)$, transition rates of $\tilde{y}$ are specified as follows:

- When a customer arrives to a station with n bikes and capacity k to retrieve a bike,

$$Q^N\left(\tilde{y}, \tilde{y} + \frac{1}{N}(\mathbf{1}_{(n-1,k)} - \mathbf{1}_{(n,k)})\right)$$

$$= \left((1-p)\lambda N \tilde{y}(n,k) + p\lambda N \frac{\tilde{y}(n,k)g(n)}{\sum_{h\in\mathcal{K}}\sum_{j=0}^{h}\tilde{y}(n,h)g(j)}\right)\mathbf{1}_{n>0}. \tag{3.25}$$

- When a customer returns a bike to a station with n bikes and capacity k,

$$Q^N\left(\tilde{y}, \tilde{y} + \frac{1}{N}(\mathbf{1}_{(n+1,k)} - \mathbf{1}_{(n,k)})\right) = \tilde{y}(n,k)N\left(\frac{M}{N} - \sum_{n'}\sum_{k'}n'\tilde{y}(n',k')\right)\mathbf{1}_{n<k}. \tag{3.26}$$

Define $\mathcal{K}^N$ as the set of unique values that station capacity k can take, then we can rewrite the ratio process R_t^N in terms of $\tilde{Y}^N(n,k)$,

$$R_t^N(j) = \sum_{k\in\mathcal{K}^N}\sum_{n=0}^{k}\tilde{Y}_t^N(n,k)\mathbf{1}\left\{\left\lfloor\frac{nK_{\max}}{k}\right\rfloor = j\right\}. \tag{3.27}$$

The benefit of this representation is that the ratio process R^N is just a linear transformation of these empirical processes $\tilde{Y}_t^N(n,k)$. Thus, we can utilize previous mean field and diffusion limit results we have shown for $\tilde{Y}_t^N(n,k)$ for fixed k, to show similar results for the ratio process. In Sections 3.1 and 3.2 we provide mean field limit and diffusion limit analysis for the ratio process.

3.1 Mean Field Limit for the Ratio Process

Now we state the main theorem in this section that proves the convergence of ratio process to its mean field limit.

Theorem 3.1 (Functional Law of Large Number). *Let $|.|$ denote the Euclidean norm in $\mathbb{R}^{\mathbb{N}\times\mathcal{K}}$. Suppose that $\lim_{N\to\infty}\frac{M}{N} = \gamma$ and $\tilde{Y}_0^N \xrightarrow{p} \tilde{y}_0$. We also assume that capacity k can only take value from a finite set, namely $\mathcal{K}$. Then we have for $\forall\epsilon > 0$*

$$\lim_{N\to\infty}P\left(\sup_{t\leq t_0}|R_t^N - r_t| > \epsilon\right) = 0$$

14

where

$$r_t(j) = \sum_{k \in \mathcal{K}} \sum_{n=0}^{k} \tilde{y}_t(n,k) \mathbf{1}\left\{ \left\lfloor \frac{n K_{\max}}{k} \right\rfloor = j \right\}$$

and $\tilde{y}_t$ is the unique solution to the following differential equation starting at $\tilde{y}_0$

$$\dot{\tilde{y}}_t = b(\tilde{y}_t) \tag{3.28}$$

where $b : [0,1]^{\mathbb{N} \times \mathcal{K}} \to \mathbb{R}^{\mathbb{N} \times \mathcal{K}}$ is a vector field satisfies

$$
\begin{aligned}
b(\tilde{y}_t) &= \sum_{k \in \mathcal{K}} \sum_{n=0}^{k} \left[\left((1-p) + p \frac{g(n)}{\sum_{j=0}^{k} \tilde{y}(j,k) g(j)} \right) (\mathbf{1}_{n-1,k} - \mathbf{1}_{n,k}) \lambda \mathbf{1}_{n>0} \right. \\
&\quad \left. + \left(\gamma - \sum_{k \in \mathcal{K}} \sum_{j=0}^{k} j \tilde{y}(j,k) \right) (\mathbf{1}_{n+1,k} - \mathbf{1}_{n,k}) \mathbf{1}_{n<k} \right] y_t(n,k) \tag{3.29}
\end{aligned}
$$

or for any $k \in \mathcal{K}$,

$$
b(\tilde{y}(0,k)) = \underbrace{-\left(\gamma - \sum_{h \in \mathcal{K}} \sum_{j=0}^{h} j \tilde{y}(j,h) \right) \tilde{y}(0,k)}_{\text{return a bike to a no-bike station}} + \underbrace{\left((1-p) + p \frac{g(1)}{\sum_{h \in \mathcal{K}} \sum_{j=0}^{K} \tilde{y}(j,h) g(j)} \right) \lambda \tilde{y}(1,k)}_{\text{retrieve a bike from a 1-bike station}},
$$

$$\tag{3.30}$$

$$
\begin{aligned}
b(\tilde{y}(i,k)) &= \underbrace{\left((1-p) + p \frac{g(i)}{\sum_{h \in \mathcal{K}} \sum_{j=0}^{K} \tilde{y}(j,h) g(j)} \right) \lambda \tilde{y}(i+1,k)}_{\text{retrieve a bike from a } i+1\text{-bike station}} + \underbrace{\left(\gamma - \sum_{h \in \mathcal{K}} \sum_{j=0}^{h} j \tilde{y}(j,h) \right) y(i-1,k)}_{\text{return a bike to a } i-1\text{-bike station}} \\
&\quad \underbrace{- \left(\left((1-p) + p \frac{g(i)}{\sum_{h \in \mathcal{K}} \sum_{j=0}^{K} \tilde{y}(j,h) g(j)} \right) \lambda + \gamma - \sum_{h \in \mathcal{K}} \sum_{j=0}^{h} j \tilde{y}(j,h) \right) \tilde{y}(i,k)}_{\text{retrieve and return a bike to a } i\text{-bike station}},
\end{aligned}
$$

$$\tag{3.31}$$

for $i = 1, ..., k-1$ and

$$
b(\tilde{y}(k,k)) = \underbrace{-\left((1-p) + p \frac{g(k)}{\sum_{h \in \mathcal{K}} \sum_{j=0}^{K} \tilde{y}(j,h) g(j))} \right) \lambda \tilde{y}(k,k)}_{\text{retrieve a bike from a } k\text{-bike station}} + \underbrace{\left(\gamma - \sum_{h \in \mathcal{K}} \sum_{j=0}^{h} j \tilde{y}(j,h) \right) \tilde{y}(k-1,k)}_{\text{return a bike to a } k-1\text{-bike station}}.
$$

$$\tag{3.32}$$

Proof. Since

$$
\begin{aligned}
|R_t^N - r_t|^2 &\leq \sum_{j=0}^{K_{max}} \left(\sum_{k \in \mathcal{K}} \sum_{n=0}^{k} (\tilde{Y}_t^N(n,k) - \tilde{y}_t(n,k)) \mathbf{1}\left\{ \left\lfloor \frac{nK_{\max}}{k} \right\rfloor = j \right\} \right)^2 \\
&\leq K_{max}(K_{max}+1) n(\mathcal{K}) \sum_{k \in \mathcal{K}} \sum_{n=0}^{k} |\tilde{Y}_t^N(n,k) - \tilde{y}_t(n,k)|^2 \\
&\leq 2K_{max}^3 \sum_{k \in \mathcal{K}} |\tilde{Y}_t^N(\cdot,k) - \tilde{y}_t(\cdot,k)|^2,
\end{aligned} \tag{3.33}
$$

for any $k \in \mathcal{K}$, by Theorem 2.3, we know that

$$
\lim_{N \to \infty} P\left(\sup_{t \leq t_0} |\tilde{Y}_t^N(\cdot,k) - \tilde{y}_t(\cdot,k)| > \epsilon \right) = 0.
$$

Thus, since $\mathcal{K}$ is a finite set, it implies

$$
\lim_{N \to \infty} P\left(\sup_{t \leq t_0} |R_t^N - r_t| > \epsilon \right) \leq \lim_{N \to \infty} P\left(\sup_{t \leq t_0} \sum_{k \in \mathcal{K}} |\tilde{Y}_t^N(\cdot,k) - \tilde{y}_t(\cdot,k)|^2 > \epsilon^2/2K_{max}^3 \right) = 0.
$$

$\square$

3.2 Diffusion Limit for the Ratio Process

We define the diffusion scaled ratio process by substracting the mean field limit from the ratio process and rescaling by $\sqrt{N}$. Thus, we obtain the following expression for the diffusion scaled ratio process

$$
Z_t^N = \sqrt{N}(R_t^N - r_t). \tag{3.34}
$$

For each $j = 0, \cdots, K_{max}$, we have

$$
\begin{aligned}
Z_t^N(j) &= \sum_{k \in \mathcal{K}} \sum_{n=0}^{k} \sqrt{N}(\tilde{Y}_t^N(n,k) - \tilde{y}_t(n,k)) \mathbf{1}\left\{ \left\lfloor \frac{nK_{\max}}{k} \right\rfloor = j \right\} \\
&= \sum_{k \in \mathcal{K}} \sum_{n=0}^{k} D_t^N(n,k) \mathbf{1}\left\{ \left\lfloor \frac{nK_{\max}}{k} \right\rfloor = j \right\}.
\end{aligned} \tag{3.35}
$$

Now we state the functional central limit theorem for the ratio process as follows.

Theorem 3.2. *Consider Z_t^N in $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^{K+1})$ with the Skorokhod J_1 topology, and suppose that $\limsup_{N \to \infty} \sqrt{N}(\frac{M}{N} - \gamma) < \infty$. Then if Z_0^N converges in distribution to Z_0, then Z_t^N converges to the unique OU process solving $Z_t = Z_0 + \int_0^t b'(y_s) Z_s ds + \tilde{M}_t$ in distribution, where*

$$
\tilde{M}_t(j) = \sum_{k \in \mathcal{K}} \sum_{n=0}^{k} M_t(n,k) \mathbf{1}\left\{ \left\lfloor \frac{nK_{\max}}{k} \right\rfloor = j \right\}
$$

for $j = 0, \cdots, K_{\max}$.

Proof. We have shown in Theorem 2.4 that

$$D_t^N(n, k) \Rightarrow D_t(n, k)$$

for any $k \in \mathcal{K}$ and $0 \leq n \leq k$. We also know that

$$\left\lfloor \frac{nK_{\max}}{k} \right\rfloor = j \Rightarrow j \leq \frac{nK_{\max}}{k} < j + 1 \Rightarrow j \cdot \frac{k}{K_{\max}} \leq n < (j+1) \cdot \frac{k}{K_{\max}}.$$

Since $k/K_{\max} < 1$, there exists at most one $n^*(j, k)$ in $\{0, \cdots, k\}$ such that $\left\lfloor \frac{n^*(j,k)K_{\max}}{k} \right\rfloor = j$. Then

$$Z_t^N(j) = \sum_{k \in \mathcal{K}} D_t^N(n^*(j, k), k). \qquad (3.36)$$

Since all k's in $\mathcal{K}$ are distinct, $\{D_t^N(n^*(j, k), k)\}_{k \in \mathcal{K}}$ are independent. This can be shown by the fact that $Y^N(n, k)$ and $Y^N(n', k')$ are independent when $k \neq k'$. Thus,

$$\sum_{k \in \mathcal{K}} D_t^N(n^*(j, k), k) \Rightarrow \sum_{k \in \mathcal{K}} D_t(n^*(j, k), k). \qquad (3.37)$$

Define

$$Z_t(j) = \sum_{k \in \mathcal{K}} D_t(n^*(j, k), k),$$

then we proved that Z_t^N converge in distribution to Z_t, and that Z_t satisfies the SDE $Z_t = Z_0 + \int_0^t b'(y_s) Z_s ds + \tilde{M}_t$. $\qquad \square$

4 Steady State Analysis

In this section, we provide the steady state analysis of the mean field limit of both the empirical measure process and the ratio process. First, we compute explicitly the equilibrium point and the Lyapunov functions for the mean field limit. We also show the interchanging of limits result for both the empirical measure process and the ratio process, that is, the stationary distribution $Y^N(\infty)$ of the empirical process converges to the same equilibrium point $\bar{y}$ of the mean field limit.

4.1 Equilibrium Analysis for Empirical Process

Our goal in this section is to prove the following diagram commutes,

$$
\begin{array}{ccc}
Y^N(t) & \xrightarrow{N \to \infty} & y(t) \\
{\scriptstyle t \to \infty} \downarrow & & \downarrow {\scriptstyle t \to \infty} \\
Y^N(\infty) & \xrightarrow[N \to \infty]{} & y(\infty)
\end{array}
$$

We have showed in Section 2.3 that $Y^N(t) \xrightarrow{p} y(t)$, we will now give the proof of the existence and uniqueness of $y(\infty)$, as well as the convergence of invariant measure $Y^N(\infty)$ to $y(\infty)$.

Proposition 4.1. *The equilibrium point $\bar{y}$ for the mean field limit exists and is unique, and it satisfies the following equations*

$$\bar{y}_n = \frac{\prod_{i=0}^{n-1} \rho_i(\bar{y})}{1 + \sum_{i=0}^{K} \prod_{k=0}^{i-1} \rho_k(\bar{y})} \tag{4.38}$$

for $n = 0, \cdots, K$ where

$$\rho_k(\bar{y}) = \frac{\gamma - \sum_{j=0}^{K} j\bar{y}_j}{\lambda \left(1 - p + p\frac{g(k)}{\sum_{j=0}^{K} g(j)\bar{y}_j}\right)} \tag{4.39}$$

for $k = 0, \cdots, K - 1$.

Proof. Denote $\bar{y}$ as the equilibrium point for the mean field limit, then this queue can be modeled as a standard birth-death process and analyzed using the basic relation

$$\left(\gamma - \sum_{j=0}^{K} j\bar{y}_j\right)\bar{y}_k = \lambda\left(1 - p + p\frac{g(k)}{\sum_{j-0}^{K} g(j)\bar{y}_j}\right)\bar{y}_{k+1}.$$

Denote

$$\rho_k(\bar{y}) = \frac{\gamma - \sum_{j=0}^{K} j\bar{y}_j}{\lambda \left(1 - p + p\frac{g(k)}{\sum_{j=0}^{K} g(j)\bar{y}_j}\right)} \tag{4.40}$$

for $k = 0, \cdots, K - 1$. Then from the recurring relationship we get

$$\bar{y}_k = \bar{y}_0 \prod_{i=0}^{k-1} \rho_i(\bar{y}).$$

Since $\sum_{k=0}^{K} \bar{y}_k = 1$, we have

$$\bar{y}_0 = \frac{1}{1 + \sum_{i=1}^{K} \prod_{k=0}^{i-1} \rho_k(\bar{y})}.$$

Then $\bar{y}$ satisfies the following equations,

$$\bar{y}_n = \frac{\prod_{i=0}^{n-1} \rho_i(\bar{y})}{1 + \sum_{i=1}^{K} \prod_{k=0}^{i-1} \rho_k(\bar{y})} \tag{4.41}$$

for $n = 0, \cdots, K$.

To show the existence and uniqueness of $\bar{y}$, we rewrite Equation (4.40) as

$$\gamma = \lambda\rho_k(\bar{y})\left(1 - p + p\frac{g(k)}{\sum_{j=0}^{K} g(j)\bar{y}_j}\right) + \sum_{j=0}^{K} j\bar{y}_j \triangleq f(\rho_k). \tag{4.42}$$

For any $0 \le k \le K - 1$, denote the right-hand side of the equation to be a function of ρ_k, then

$$f'(\rho_k) = \lambda(1-p) + \lambda pg(k)\frac{\sum_{j=0}^{K} g(j)\bar{y}_j - \rho_k \sum_{j=0}^{K} g(j)\bar{y}'_j(\rho_k)}{\left(\sum_{j=0}^{K} g(j)\bar{y}_j\right)^2} + \sum_{j=0}^{K} j\bar{y}'_j(\rho_k). \tag{4.43}$$

18

Since

$$\bar{y}_j - \rho_k \bar{y}'_j(\rho_k) = \frac{\prod_{i=0}^{j-1} \rho_i}{1 + \sum_{i=0}^{K-1} \prod_{h=0}^{i} \rho_h} - \frac{\prod_{i=0}^{j-1} \rho_i \left(\mathbf{1}_{j>k} - \sum_{i=k}^{K-1} \prod_{h=0}^{i} \rho_h\right)}{\left(1 + \sum_{i=0}^{K-1} \prod_{h=0}^{i} \rho_h\right)^2}$$

$$= \frac{\prod_{i=0}^{j-1} \rho_i}{\left(1 + \sum_{i=0}^{K-1} \prod_{h=0}^{i} \rho_h\right)^2} \left(1 - \mathbf{1}_{j>k} + \sum_{i=k}^{K-1} \prod_{h=0}^{i} \rho_h\right)$$

$$\geq 0, \tag{4.44}$$

the second term of $f'(\rho_k)$ is positive. We also have that

$$\bar{y}'_j(\rho_k) < 0 \text{ when } j \leq k$$

$$\bar{y}'_j(\rho_k) > 0 \text{ when } j > k$$

which means the third term $\sum_{j=0}^{K} j \bar{y}'_j(\rho_k)$ is also positive. Thus, $f(\rho_k)$ is an increasing function of ρ_k on $\mathbb{R}^+$ and Equation (4.42) has a unique root $\rho_k > 0$. Therefore, we have shown that $\bar{y}$ exists and is unique. $\qquad\square$

Now we proceed to showing a Lyapunov function of the mean field limit $y(t)$. To simplify the notation, we denote the distribution $\nu_{\rho(y)}$ as

$$\nu_{\rho(y)}(n) = \frac{\prod_{i=0}^{n-1} \rho_i(y)}{1 + \sum_{i=1}^{K} \prod_{k=0}^{i-1} \rho_k(y)}$$

for $n = 0, \cdots, K$ where

$$\rho_k(y) = \frac{\gamma - \sum_{j=0}^{K} j y_j}{\lambda \left(1 - p + p \frac{g(k)}{\sum_{j=0}^{K} g(j) y_j}\right)}$$

for $k = 0, \cdots, K - 1$. We also define

$$Z(\rho) = 1 + \sum_{i=1}^{K} \prod_{k=0}^{i-1} \rho_k$$

and

$$\phi_{k,n}(y_n) = \rho_k(y)$$

Then we have the following theorem of the Lyapunov function for $y(t)$.

Theorem 4.2 (Lyapunov function). *Denote $\mathcal{X} = \{0, 1, \cdots, K\}$ and $\mathcal{P}(\mathcal{X})$ as the set of all probability measures on $\mathcal{X}$. Then for any $y \in \mathcal{P}(\mathcal{X})^\circ$, define $g(y) = h(y) - \log(Z(\rho(y))) + \sum_{n=0}^{K} \sum_{k=0}^{n-1} S_{k,n}(y_n))$ where*

$$h(y) = \sum_{n=0}^{K} y_n \log(y_n / \nu_{\rho(y)}(n))$$

is the relative entropy of y w.r.t $\nu_{\rho(y)}$, and $S_{k,n}$ is some primitive of $x \mapsto x \frac{\phi'_{k,n}(x)}{\phi_{k,n}(x)}$. Then g is a Lyapunov function for the dynamical system $y(t)$.

19

Proof. The proof is provided in the Appendix (Section 7). □

Now we state the main theorem of this section, the interchanging of limits result for the mean field limit.

Theorem 4.3. *The sequence of invariant probability distributions $\bar{y}^N$ of process Y_t^N converges weakly to he Dirac mass at $\bar{y}$ as $N \to \infty$.*

Proof. The Lyapunov function we introduced in Theorem 4.2 plays a key role in the proof of this result. We will be showing the result in the following steps:

1) The infinitesimal generator Q_N of the empirical measure process converges to the degenerate generator $Q_\infty(\cdot)(y)$, which satisfies $Q_\infty(F)(y) = \langle \nabla F(y), b(y) \rangle$ for any $y \in \mathcal{P}(\mathcal{X})$ and $F \in C^2(\mathbb{R}^{K+1})$.

2) For any subsequence $\bar{y}^{N_p}$ of the invariant measures $\bar{y}^N$, its limit $\tilde{y}$ solves $Q_\infty \tilde{y} = 0$, which means it is the invariant measure of the markov chain with infinitesimal generator Q_∞.

3) Using the Lyapunov function g, we can show that $Q_\infty(g)\tilde{y} = \int_{\mathcal{P}(\mathcal{X})} y B_y \nabla g(y) \tilde{y}(dy) = 0$. The integral being non-positive, and zero only at $\bar{y}$, means $\tilde{y}$ is also the equilibrium of the dynamical system.

4) By uniqueness of the equilibrium point and compactness of $\mathcal{P}(\mathcal{X})$, $\bar{y}$ is the only weak limit of $\bar{y}^N$.

Now define the functional operator associated with the Q-matrix as

$$
\begin{aligned}
Q_N(F)(y) &= \sum_{z \neq y} Q_N(y, z)(F(y) - F(z)) \\
&= \sum_{n=0}^{K} \left[\left((1-p) + p\frac{g(n)}{\sum_{j=0}^{K} y_j g(j)} \right) \left(F\left(y + \frac{1}{N}(\mathbf{1}_{n-1} - \mathbf{1}_n)\right) - F(y) \right) \lambda N \mathbf{1}_{n>0} \right. \\
&\quad \left. + N\left(\frac{M}{N} - \sum_{j=0}^{K} j y_j\right) \left(F\left(y + \frac{1}{N}(\mathbf{1}_{n+1} - \mathbf{1}_n)\right) - F(y) \right) \mathbf{1}_{n<K} \right] y_n. \quad (4.45)
\end{aligned}
$$

for any $F \in C^2(\mathbb{R}^{K+1})$. Now, it is easy to check that the sequence $Q_N(F)(y)$ converges to the following

$$
\begin{aligned}
\sum_{n=0}^{K} &\left[\left((1-p) + p\frac{g(n)}{\sum_{j=0}^{K} y_j g(j)} \right) \langle \nabla F(y), \mathbf{1}_{n-1} - \mathbf{1}_n \rangle \lambda \mathbf{1}_{n>0} \right. \\
&\quad \left. + \left(\gamma - \sum_{j=0}^{K} j y_j \right) \langle \nabla F(y), \mathbf{1}_{n-1} - \mathbf{1}_n \rangle \mathbf{1}_{n<K} \right] y_n. \quad (4.46)
\end{aligned}
$$

which we define as $Q_\infty(F)(y)$. Moreover, by using Taylor's formula at the second order, this convergence is uniform with respect to $y \in \mathcal{P}(\mathcal{X})$. Thus we have

$$
Q_\infty(F)(y) = \langle \nabla F(y), b(y) \rangle, y \in \mathcal{P}(\mathcal{X}).
$$

20

Now since $\mathcal{X}$ is a finite set, the set of $\mathcal{P}(\mathcal{X})$ is compact, and the sequence of invariant measure $\bar{y}^N$ is relatively compact. Let $\tilde{y}$ be the limit of some subsequence $\bar{y}^{N_p}$. Then for any $F \in C^2(\mathbb{R}^{K+1})$ and $p \geq 0$,

$$\int_{\mathcal{P}(\mathcal{X})} Q_{N_p}(F)(y)\bar{y}^{N_p}(dy) = 0.$$

The uniform convergence of $Q_{N_p}(F)$ to $Q_\infty(F)$ implies

$$0 = \int_{\mathcal{P}(\mathcal{X})} Q_\infty(F)(y)\tilde{y}(dy) = \int_{\mathcal{P}(\mathcal{X})} \langle \nabla F(y), b(y) \rangle\, \tilde{y}(dy).$$

Then $\tilde{y}$ is an invariant measure of the Markov process associated with the infinitesimal generator Q_∞.

Now we show that $y(t)$ must lies in the interior of $\mathcal{P}(\mathcal{X})$. For $t \geq 0$, denote $y(x, t)$ as the dynamical system at time t starting from $y_0 = x \in \mathcal{P}(\mathcal{X})$. Assume that there exists $x \in \mathcal{P}(\mathcal{X})$, $s > 0$ and $n \in \mathcal{X}$ such that $y_n(x, s) = 0$. Since $y_n(x, t)$ is non-negative and the function $t \to y(x, t)$ is of class C^1, it implies that $b_n(y(x, s)) = \dot{y}_n(x, s) = 0$. By the definition of $b_n(y)$ and that $y_n(x, s) = 0$, we know that $y_{n+1}(x, s) = y_{n-1}(x, s) = 0$ when $n+1$ or $n-1$ is in $\mathcal{X}$. By repeating the same argument, we can conclude that all coordinates of $y(x, s)$ are zeros. This contradicts the fact that $y(x, s)$ is a probability measure. Hence the boundary $\partial\mathcal{P}(\mathcal{X})$ can't be reached by $y(x, t)$ in positive time.

For $x \in \mathcal{P}(\mathcal{X})$ and $0 < s' < s$, since $g(y(x, \cdot))$ is in class C^1 on $[s', s]$ and its derivative is $\langle \nabla g(y(x, \cdot)), b(y(x, \cdot)) \rangle$, we have

$$g(y(x, s)) - g(y(x, s')) = \int_{s'}^{s} Q_\infty(g)(y(x, u))du.$$

By integrating both sides w.r.t $\tilde{y}$, the invariance of $\tilde{y}$ for the process $y(x, t)$ and Fubini's theorem show that

$$\begin{aligned}
0 &= \int_{\mathcal{P}(\mathcal{X})} g(y(x, s))\tilde{y}(dx) - \int_{\mathcal{P}(\mathcal{X})} g(y(x, s'))\tilde{y}(dx) \\
&= \int_{\mathcal{P}(\mathcal{X})} \int_{s'}^{s} Q_\infty(g)(y(x, u))du\tilde{y}(dx) \\
&= (s - s') \int_{\mathcal{P}(\mathcal{X})} Q_\infty(g)(x)\tilde{y}(dx).
\end{aligned} \tag{4.47}$$

By Theorem 4.2, $\langle \nabla g(x), b(x) \rangle \leq 0$ for all x, thus $Q_\infty(g)(x) = 0$ $\tilde{y}$-almost surely. Thus $\tilde{y}$ is also the equilibrium of the dynamical system. Since we showed in Proposition 4.1 that the equilibrium point $\bar{y}$ is unique, we can conclude that the sequence of invariant probability measure $\bar{y}^N$ of process $Y^N(t)$ converges weakly to he Dirac mass at $\bar{y}$ as $N \to \infty$.

$\square$

4.2 Equilibrium Analysis for Ratio Process

In this section, we provide a similar equilibrium analysis for the ratio process R_t^N, which utilize the results we obtained in Section 4.1.

Proposition 4.4. *The equilibrium point $\bar{r}$ for the mean field limit of the ratio process $r(t)$ exists and is unique, and it satisfies the following equation*

$$\bar{r} = \sum_{k \in \mathcal{K}} y_K(k)\bar{y}(\cdot|K = k)A^{(k)}$$

where for each $k \in \mathcal{K}$, $A^{(k)}$ is a $(k+1) \times (K_{\max} + 1)$ matrix defined as

$$A_{ij}^{(k)} = \mathbf{1}\left\{ \frac{iK_{\max}}{k} = j \right\}$$

for $i = 0, \cdots, k$ and $j = 0, \cdots, K_{\max}$ and

$$y_K(k) = \sum_{n=0}^{k} \tilde{y}(n, k)$$

is the marginal distribution for capacity k.

Proof. By the definition of the ratio process in Section 3, the ratio process can be written as

$$R_t^N = \sum_{k \in \mathcal{K}^N} \sum_{n=0}^{k} \tilde{Y}_t^N(n, k)\mathbf{1}\left\{ \left\lfloor \frac{nK_{\max}}{k} \right\rfloor = j \right\} = \sum_{k \in \mathcal{K}^N} \tilde{Y}_t^N(\cdot, k)A^{(k)}.$$

Similarly, the mean field limit for the ratio process R_t^N can be written as

$$r_t = \sum_{k \in \mathcal{K}} \tilde{y}_t(\cdot, k)A^{(k)}.$$

Let the marginal distribution for capacity k be $y_K(k) = \sum_{n=0}^{k} \tilde{y}(n, k)$, then the network of all stations with capacity k can be viewed as a closed queueing network and its empirical mean field limit $y_t(\cdot|K = k)$ can be written as

$$y_t(n|K = k) = \frac{\tilde{y}_t(n, k)}{y_K(k)}$$

for $n = 0, \cdots, k$. We have shown the uniqueness and existence of equilibrium points for queueing networks with same capacities, therefore the equilibrium points $\bar{r}$ for the ratio process also exists and is unique, and it satisfies

$$\bar{r} = \sum_{k \in \mathcal{K}} y_K(k)\bar{y}(\cdot|K = k)A^{(k)}.$$

$\square$

Theorem 4.5. *The sequence of invariant probability distributions $\bar{r}^N$ of the ratio process R_t^N converges weakly to the Dirac mass at $\bar{r}$ as $N \to \infty$.*

Proof. By Theorem 4.3, the invariant probability distribution $\bar{y}^N$ of process $Y^N(t)$ converges weakly to the Dirac mass at $\bar{y}$ as $N \to \infty$. We know that the invariant probability distribution $\bar{r}^N$ of ratio process R_t^N can be expressed as

$$\bar{r}^N(j) = \sum_{k \in \mathcal{K}} \sum_{n=0}^{k} \bar{y}^N(n,k) \mathbf{1} \left\{ \left\lfloor \frac{nK_{\max}}{k} \right\rfloor = j \right\},$$

and that

$$\bar{r}(j) = \sum_{k \in \mathcal{K}} \sum_{n=0}^{k} \bar{y}(n,k) \mathbf{1} \left\{ \left\lfloor \frac{nK_{\max}}{k} \right\rfloor = j \right\}$$

for $j = 0, \cdots, K_{\max}$.

We also know that

$$\left\lfloor \frac{nK_{\max}}{k} \right\rfloor = j \Rightarrow j \leq \frac{nK_{\max}}{k} < j+1 \Rightarrow j \cdot \frac{k}{K_{\max}} \leq n < (j+1) \cdot \frac{k}{K_{\max}}.$$

Since $k/K_{\max} < 1$, there exists at most one $n^*(j,k)$ in $\{0, \cdots, k\}$ such that $\left\lfloor \frac{n^*(j,k)K_{\max}}{k} \right\rfloor = j$. Then

$$\bar{r}^N(j) \;=\; \sum_{k \in \mathcal{K}} \bar{y}^N(n^*(j,k), k). \tag{4.48}$$

Since convergence in distribution to a constant implies convergence in probability, and by additivity of convergence in probability, we have

$$\sum_{k \in \mathcal{K}} \bar{y}^N(n^*(j,k), k) \xrightarrow{p} \sum_{k \in \mathcal{K}} \bar{y}(n^*(j,k), k), \tag{4.49}$$

which implies

$$\sum_{k \in \mathcal{K}} \bar{y}^N(n^*(j,k), k) \Rightarrow \sum_{k \in \mathcal{K}} \bar{y}(n^*(j,k), k). \tag{4.50}$$

This shows that the invariant measure $\bar{r}^N$ converges weakly to the Dirac mass at $\bar{r}$ as $N \to \infty$. $\qquad \square$

5 Insights from Numerical Examples and Simulation

In this section, we show the different aspects of the impact of information on the empirical measure. This includes how the proportion of users that have information (parameter p), different choice function g (exponential, minimum choice functions) and users' sensitivity to information (parameters in g) affect the behavior of empirical measure over time. We divide the section into two subsections, where Section 5.1 considers empirical measure under stationary arrivals, and Section 5.2 considers empirical measure under non-stationary arrivals.

5.1 Stationary Arrivals

In this section, we consider the bike sharing system under stationary arrivals. Common parameter values used in all numerical examples are $\lambda = 1, \mu = 1, K = 20, \gamma = 10$. We explore the impact of information on three different perspectives: parameter p, different choice function g, and parameters in each choice function g.

5.1.1 Impact of proportion of users with information (p) to the empirical measure

We first explore impact of parameter p, the proportion of user with information, on the empirical measure. Here we consider three different choice function: 1) exponential choice function $g(x) = e^{\theta x}$; 2) minimum choice function: $g(x) = \min(x, c)$; 3) polynomial choice function: $g(x) = x^\alpha$.

Another aspect we considered is the entropy of the empirical process. We define the **entropy** of the empirical measure as

$$S^N(t) = -\sum_{i=0}^{K} Y_t^N(i) \log(Y_t^N(i)). \tag{5.51}$$

In our model, our goal is to minimize the proportion of empty and full stations, thus we want the empirical measure to be close to a mass at the average number of bikes per stations, that is, low entropy.

Exponential Choice Function $g(x) = e^{\theta x}$ We first explore the exponential choice function. This represent a high user sensitivity to information about bike availability, as the probability of choosing stations with more bikes grows exponentially.

Figure 3 shows the empirical process at equilibrium for $p = 0, 0.25, 0.5, 1$ and $\theta = 1$. We observe that with no information ($p = 0$) the distribution looks relatively flat. As you increase the value of p, i.e. the proportion of users with information, the distribution narrows down to be centered around its mean quite drastically. We also observe that the proportion of empty and full stations go close to 0 with only 25% of user having information.

Figure 4 shows the surface plots of the mean field limit on the $p - \theta$ plane with $p \in [0, 1]$ and $\theta \in [0, 2]$. Particularly, we look at proportion of stations with 0,1,19 and 20 bikes. These plots show that proportion of nearly empty and full stations go down dramatically both when p increase from 0 to 1 and θ increase from 0 to 2.

Figure 5 shows the surface plot of the entropy of the mean field limit on the $p - \theta$ plane. We observe that the entropy goes down as p, θ increase, which is consistent with the findings in Figures 4 and 5.
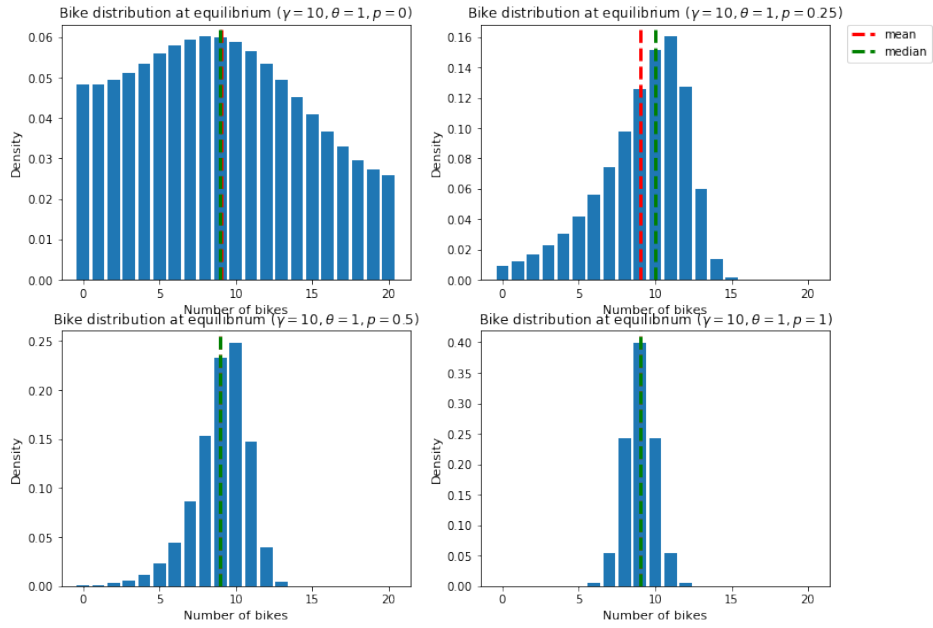
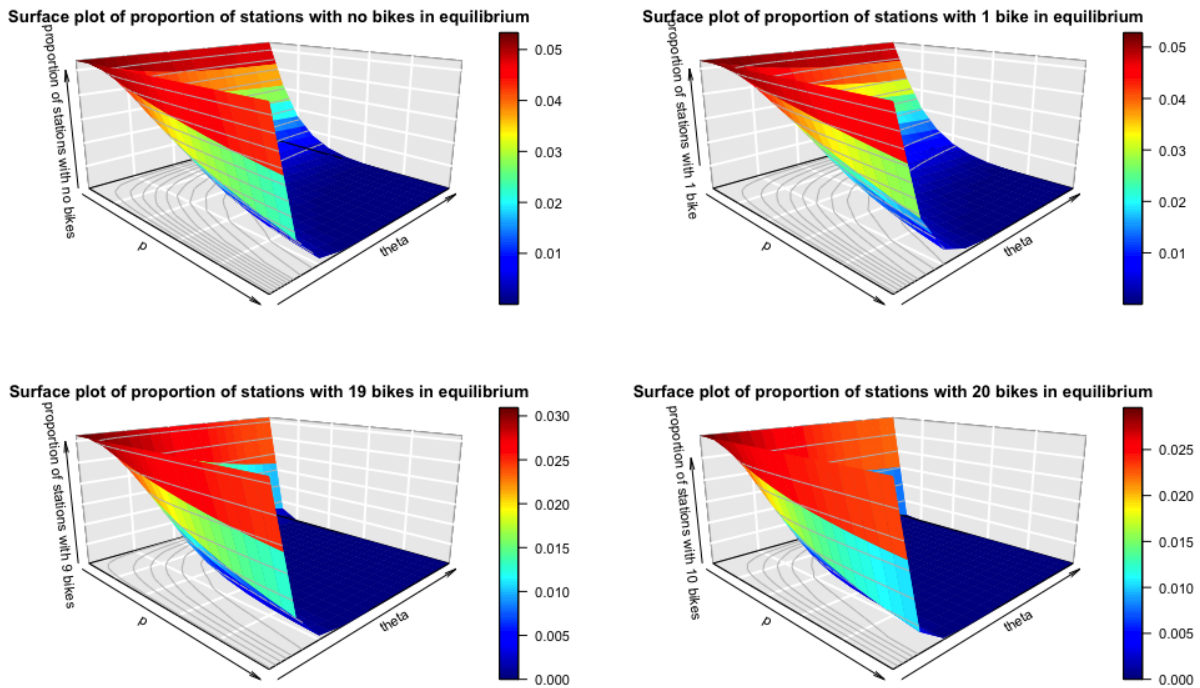Figure 3: Bike distribution for different values of p when $\theta = 2$

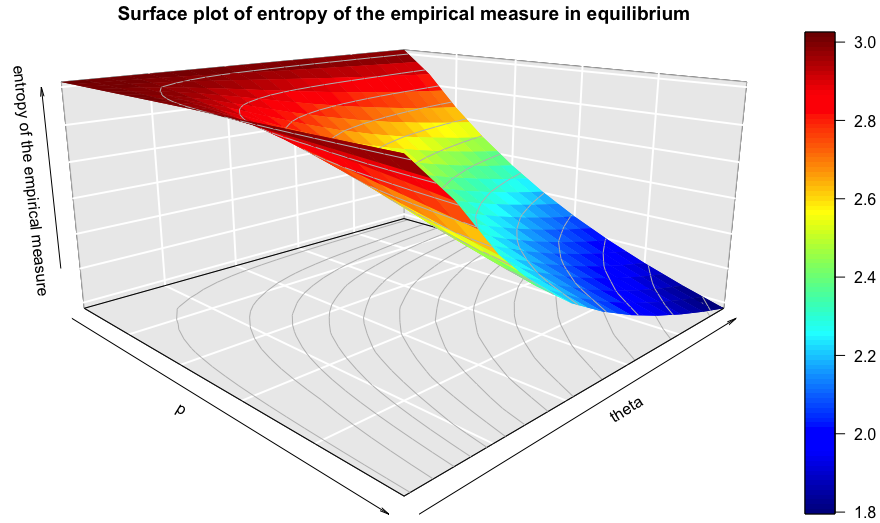Figure 4: Surface plot of $y_\infty(p, \theta)$ on the $p - \theta$ plane

Figure 5: Surface plot of the entropy of the empirical measure in equilibrium on the $p - \theta$ plane

Minimum Choice Function $g(x) = \min(x, c)$ We then explore the minimum choice function. This choice function represents that user sensitivity to information about bike availability has a threshold, as the probability of choosing stations with more bikes grows linearly but then remain the same once the threshold c is reached.

Figure 6 shows the empirical process at equilibrium for $p = 0, 0.25, 0.5, 1$ and $c = 5$. We observe that with no information ($p = 0$) the distribution looks relatively flat. As the value of p increase to 1, i.e. the proportion of users with information, the proportion of stations with empty stations goes down to close to 0. However, we don't see much changes in terms of the proportion of full or nearly full stations. This agrees with the intuition that when there is a threshold on the incentives the number of bikes available has on users, its effect on reducing full stations is diminished.

Figure 7 shows the plots of the mean field limit vs. proportion of users with information (p) for different values of c. Particularly, we look at proportion of stations with 0,1,19 and 20 bikes. These plots show that proportion of empty stations go down both when p increase from 0 to 1 and c increase from 1 to 5. However, the proportion of full stations doesn't change much for different values of p and c.

Figure 8 shows the plot of the entropy of the mean field limit vs. proportion of users with information (p) for different values of c. We observe that the entropy goes down as p, c increase, which is consistent with the findings in Figures 7 and 8.
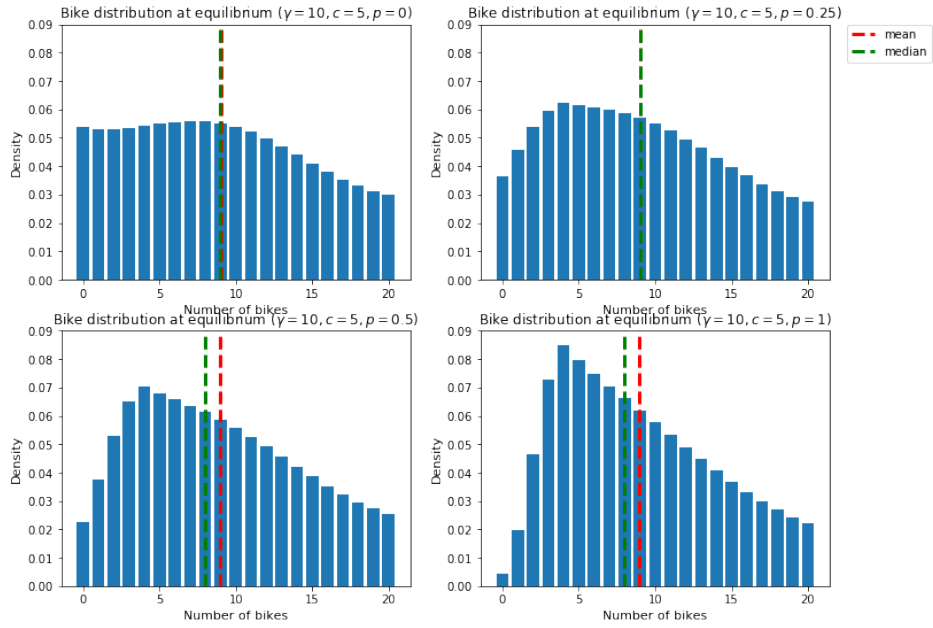
Figure 6: Bike distribution for different values of p when $K = 20$, $\gamma = 10$ and $c = 5$
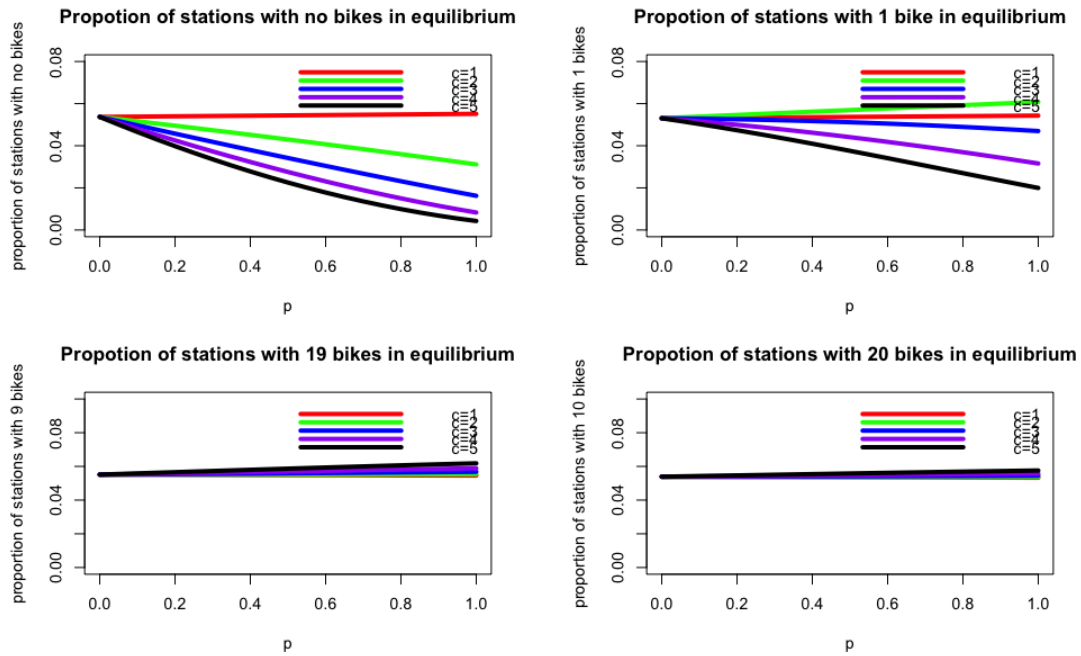

Figure 7: Proportion of stations with no bikes, 1 bike, 19 bikes and 20 bikes for different value of c
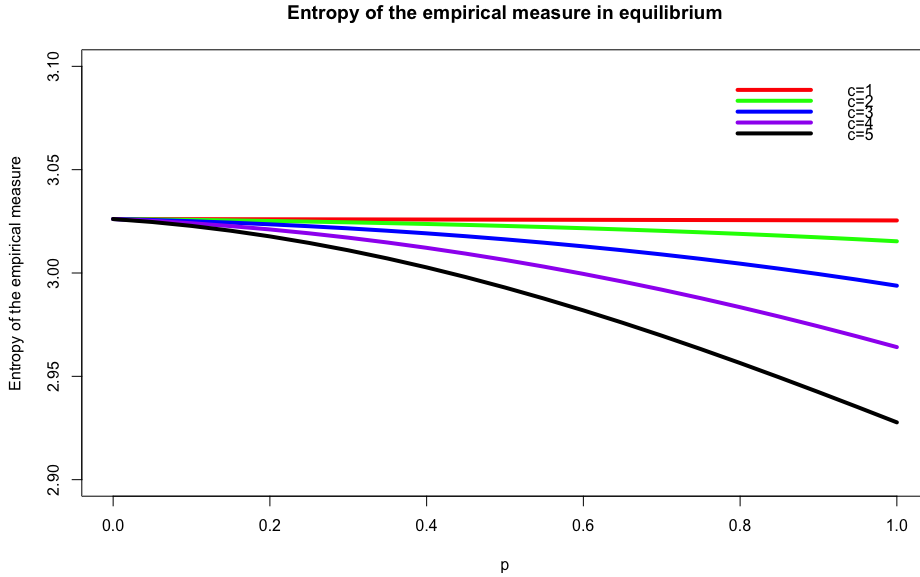
Figure 8: Entropy of the empirical measure in equilibrium for different value of c

Polynomial Choice Function $g(x) = x^\alpha$ Finally we then explore the polynomial choice function. This choice function represents a user sensitivity to information about bike availability that is in between the exponential choice function and the minimum choice function, as the probability of choosing stations with more bikes has a polynomial growth.

Figure 9 shows the empirical process at equilibrium for $p = 0, 0.25, 0.5, 1$ and $\alpha = 2$. We observe that with no information ($p = 0$) the distribution looks relatively flat. As the value of p increase to 1, i.e. the proportion of users with information, the proportion of empty and full stations goes down to close to 0. The effect of polynomial choice function is similar to that of exponential choice function, but less extreme.

Figure 10 shows the surface plots of the mean field limit on the $p - \alpha$ plane with $p \in [0, 1]$ and $\alpha \in [0, 5]$. Particularly, we look at proportion of stations with 0,1,19 and 20 bikes. These plots show that proportion of nearly empty and full stations go down dramatically both when p increase from 0 to 1 and α increase from 0 to 5. Its effect is similar to those from the exponential choice function.

Figure 11 shows the surface plot of the entropy of the mean field limit on the $p - \alpha$ plane. We observe that the entropy goes down as p, θ increase, which is consistent with the findings in Figures 10 and 11.
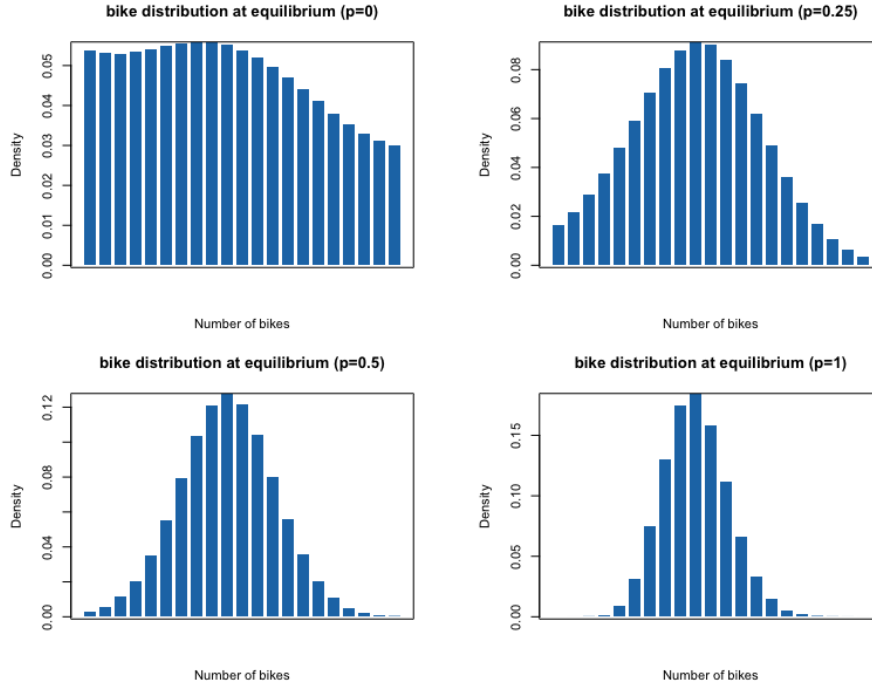
Figure 9: Bike distribution for different values of p when $K = 20$, $\gamma = 10$ and $\alpha = 2$
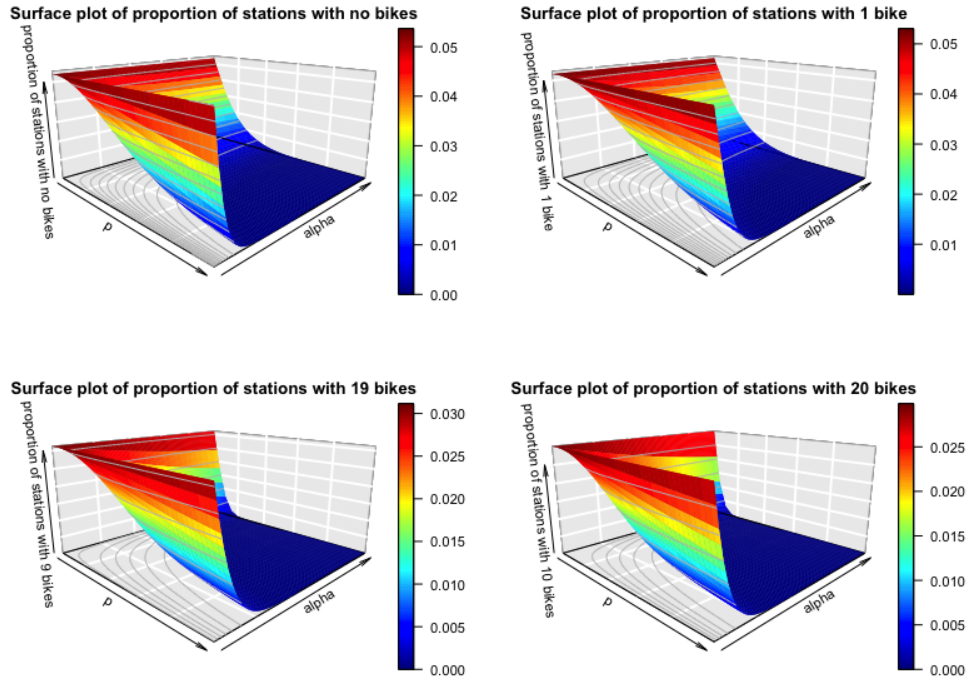

Figure 10: Surface plot of $y_\infty(p, \alpha)$ on the $p - \alpha$ plane when $K = 20$
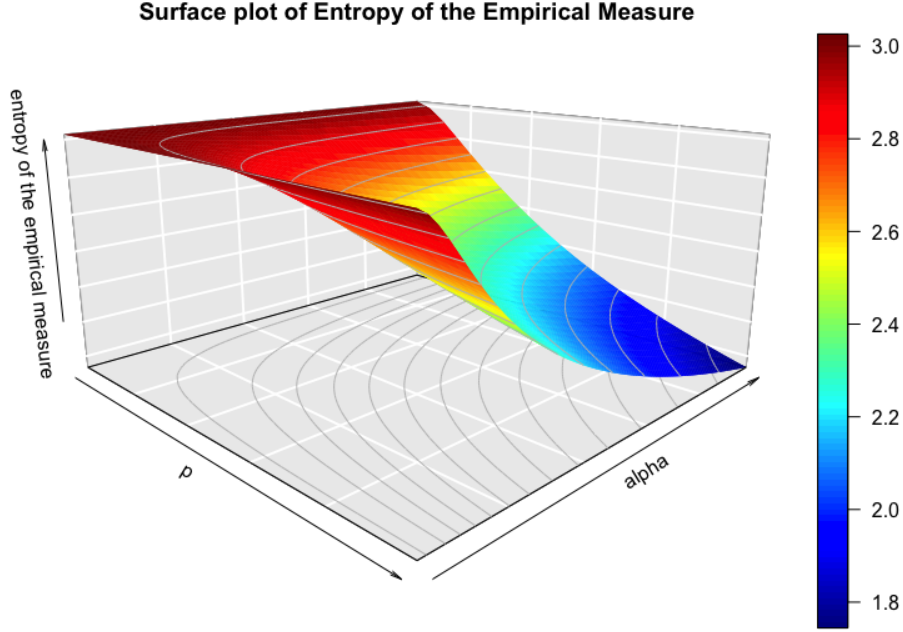
Surface plot of Entropy of the Empirical Measure

Figure 11: Surface plot of the entropy of the empirical measure in equilibrium on the $p - \alpha$ plane when $K = 20$

5.2 Non-stationary Arrivals

In this section, we consider the bike sharing system under non-stationary arrivals and provide numerical examples to analyze the effect of choice modeling. This is important because in real life bike sharing systems, arrival rates are often non-stationary and periodic.

5.2.1 Fitting Arrival Rate Function

Given the periodic property of the arrival rate function we observe from the historical data of CitiBike (see Figure 12), we perform linear regression using the following Fourier basis,

$$\lambda(t) = \sum_{i=0}^{n} \beta_{1,j} \sin(2\pi t[j/\omega]) + \beta_{2,j} \cos(2\pi t[j/\omega]), \tag{5.52}$$

where the unit of t is hour and $\omega = 24$ is the period of $\lambda(t)$.

Since the arrival rate patterns is quite different between weekdays and weekends, we will fit two separate regression on them.

Weekdays The fitted result of $\lambda(t)$ during the weekdays for $n = 5$ is the following,

$$
\begin{aligned}
\lambda(t) \quad = \quad & -43.4 \sin(2\pi t/24) - 49.5 \cos(2\pi t/24) - 38.2 \sin(4\pi t/24) - 40.0 \cos(4\pi t/24) \\
& +30.1 \sin(6\pi t/24) + 23.7 \cos(6\pi t/24) + 14.6 \sin(8\pi t/24) - 1.4 \cos(8\pi t/24) \\
& -29.4 \sin(10\pi t/24) + 1.4 \cos(10\pi t/24) + 91.4
\end{aligned}
\tag{5.53}
$$

Table 2 shows the R^2 statistics for different choice of n. We can see that as we increase n, which indicates the complexity of function form, we get better approximation to the actual arrival rate function. With $n = 5$ we obtain an R^2 statistics of 0.947, while the benefit of adding more complexity diminishes as n increase from 5 to 10.

n	1	2	3	4	5	6	7	8	9	10
R^2 statistics	0.41	0.70	0.84	0.86	0.947	0.955	0.962	0.972	0.973	0.975

Table 2: R^2 statistics for different degree n (weekdays)

Figure 12 shows the actual vs. predicted arrival rate of customers (average number of trips during each 5 minutes) on weekdays from CitiBike data. We can see that the fitted results are very accurate in comparison to the real data.

Weekends The fitted result of $\lambda(t)$ during the weekends for $n = 2$ is the following,

$$\hat{\lambda}(t) \;\; = \;\; -43.8\sin(2\pi t/24) - 39.5\cos(2\pi t/24) + 6.0\sin(4\pi t/24) + 6.7\cos(4\pi t/24) + 58.6. \tag{5.54}$$

Table 3 shows the R^2 statistics for different degree n. Since the arrival rate pattern is much simpler during weekends than in weekdays, the approximation is much easier. With $n = 1$ we already obtain an R^2 statistics of 0.955.

n	1	2	3
R^2 statistics	0.955	0.977	0.987

Table 3: R^2 statistics for different degree n (weekends)
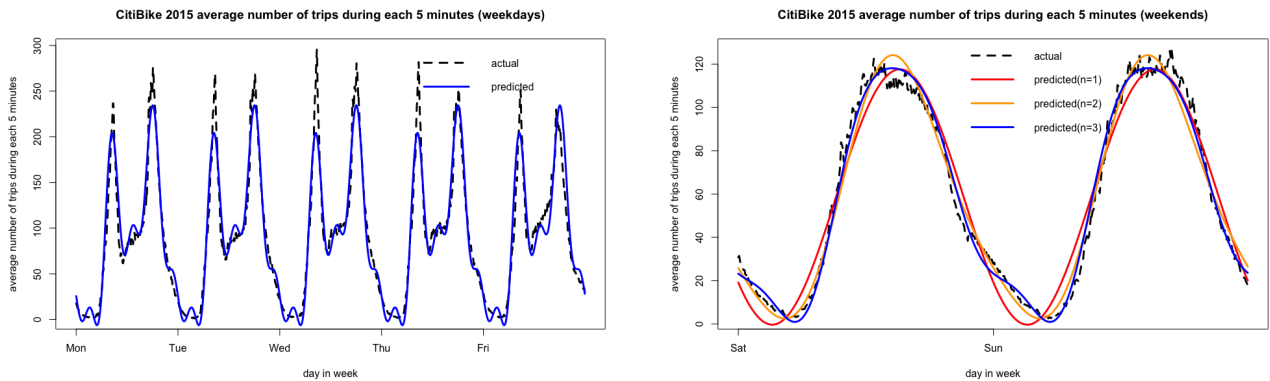

Figure 12: Fitted Arrival Rate Function (weekdays (left) and weekends (right))

Figure 12 shows the actual vs. predicted arrival rate of customers (average number of trips during each 5 minutes) on weekends from CitiBike data. Again, we can see that the fitted results are very accurate in comparison to the real data for just $n = 2$.

5.2.2 Impact of giving customer information (p) to the empirical measure

In the case of non-stationary arrivals, a figure is not enough to visualize the changes of bike distribution over time for us to analyze the impact of choice modeling. Thus, we create videos of the empirical measure process over time below.

We provide a video in this link the empirical measure over time under non-stationary arrival rates with exponential choice function for $p = 0, 0.25, 0.5, 1$. The arrival rates functions used in the video are from Equations (5.53) and (5.54) from Section 5.2.1. At time 0, the distribution is initialized at a mass at 10 bikes per station. We observe from the video that at any given time, as p, the proportion of users with information, increase from 0 to 1, the proportion of empty and full stations decrease dramatically. This is consistent with our finding in Section 5.1. We also observe that, the empirical process where p is larger is more centered around its mean, i.e. more stable through out the length of simulation, and the effect is getting bigger when p increase from 0 to 1. This suggests that with the exponential choice function, more people having information about the system will make the system more stable (in terms of less empty/full stations).

We provide a video in this link the empirical measure over time under non-stationary arrival rates with minimum choice function for $p = 0, 0.25, 0.5, 1$. The arrival rates function and the initial setting is the same as in previous video with exponential choice function. We observe that at any given time, as p increase from 0 to 1, the proportion of empty stations decrease dramatically. However there is very little impact on full or nearly full stations. This is consistent with our finding in Section 5.1. We also observe that the empirical process with larger p has less empty or nearly empty stations throughout the duration of the simulation, not not much difference in terms of full or nearly full stations. This suggests that the minimum choice function, more people having information about the system will give a better control on the proportion of empty or nearly empty stations, however it has very little impact on full or nearly full stations.

Now we look at a simpler example, where $K = 3$ and $\lambda(t) = 1 + 0.5\sin(t/2)$, and examine each individual component of the empirical process. That is, we show the surface plots of the fluid limit of the proportion of stations with 0,1,2,3 bikes individually on the time-p plane under non-stationary arrivals. Here we use the exponential choice function as an example. Parameters are given as $\theta = 2$ and initial distribution of bikes is $y(0) = (0.25, 0.25, 0.25, 0.25)$.

In Figure 13, we can see that as p increases, the average proportion of empty stations goes down. However, the amplitude of the variation of the proportion of empty stations becomes larger, which means the system is more unstable. In terms of full stations, we see the opposite: both the average proportion of full stations and the amplitude of variation go down. This tells us that in the time-varying arrival case, we may not want to give customer too much information.

In Figure 14, we show the surface plot of the entropy of the empirical process on the time-p plane. We observe that the average entropy is decreasing as p increase and we do not see huge changes in amplitude of the variation of entropy over time.

In summary, this example suggests that under the non-stationary arrivals setting, it might not always be better to give users more information. Whether giving people information will lower proportion of empty or full stations depends on the specific parameters setting, and policy makers of bike sharing companies should be aware of this. However, it is very easy to

verify the benefit of giving users information using the mean field limit and diffusion limit results to give approximations to the mean and variance of empirical measure process under non-stationary setting.
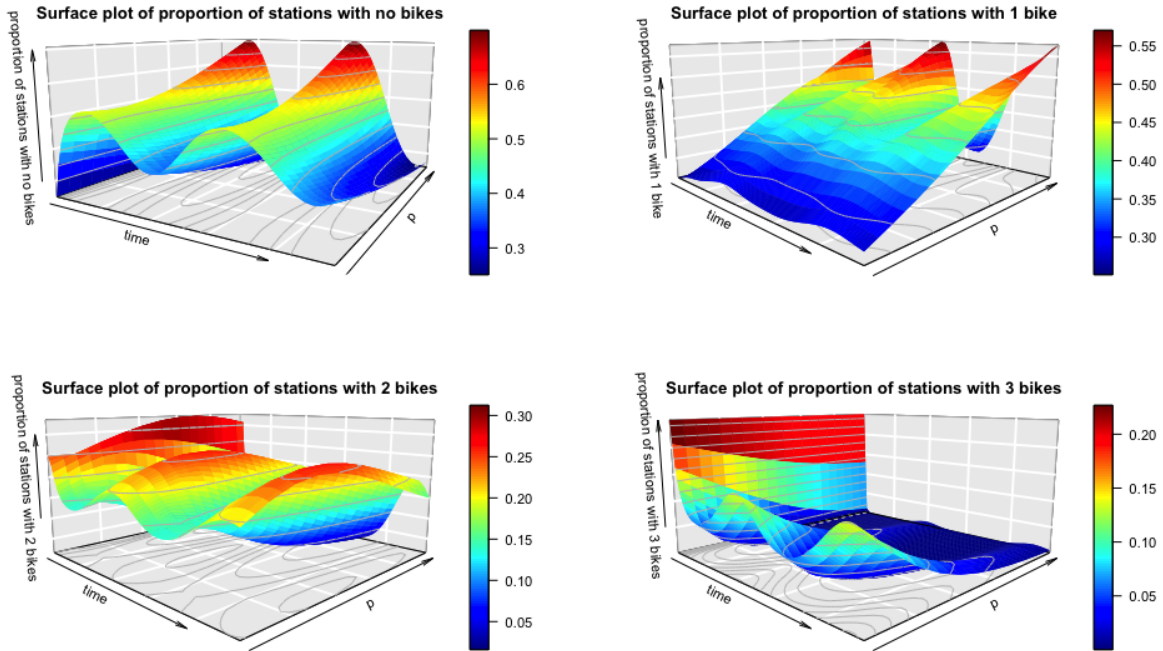

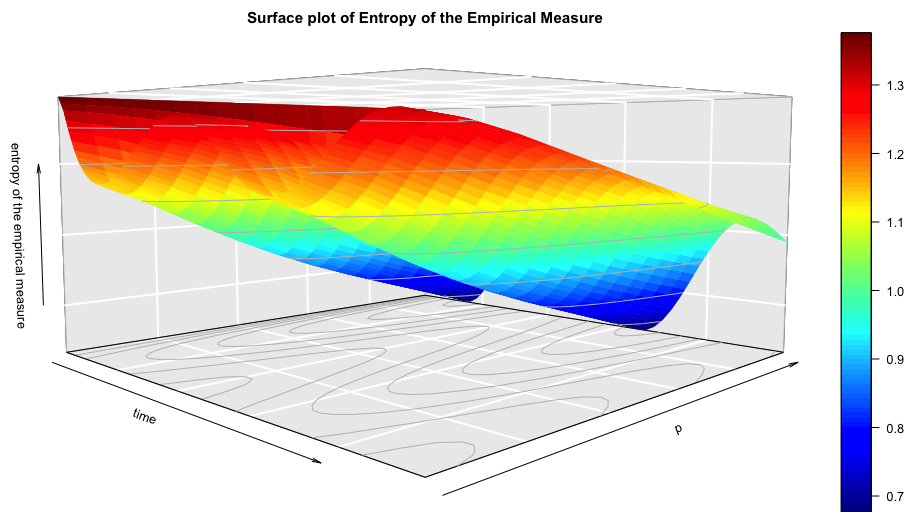Figure 13: Surface plot of $y(t,p)$ on the time-p plane when $K = 3$

Figure 14: Surface plot of entropy of the empirical measure on the time-p plane when $K = 3$

5.2.3 Impact of user's sensitivity to information (θ) to the empirical measure

In this section, we explore the impact of user's sensitivity to information on the empirical measure process. Specifically, we consider the exponential choice function.

In Figure 15 we show the surface plots of the fluid limit on the time-θ plane, with $K = 3$, $p = 0.5$ and $\theta \in [0, 5]$. The initial distribution of bikes is $y(0) = (0.25, 0.25, 0.25, 0.25)$. We observe that the the average proportion of empty stations slightly goes down, but the amplitude of the variation becomes larger. This suggests that as users become more sensitive to information, there exists a trade-off between system average and system variation under non-stationary setting in terms of controlling empty stations. Depending on the goal, it might not always be better when users are more sensitive to information.

In Figure 16 we show the surface plot of the entropy of the empirical measure process on the time-θ plane. We observe that the average entropy and the amplitude of variation of the entropy is not monotone on θ, indicating that there exists an optimal value of θ that minimize the mean (variance) of the entropy under the non-stationary setting.

In summary, this example suggests that under the non-stationary arrivals setting, it might not always be better when customers are more sensitive to information. Whether people being more sensitive to information will lower proportion of empty or full stations depends on the specific parameters setting. There might be a trade-off between the system average performance and the system variation, in which case an optimal sensitivity parameter θ exists where the empirical measure achieves minimal entropy.
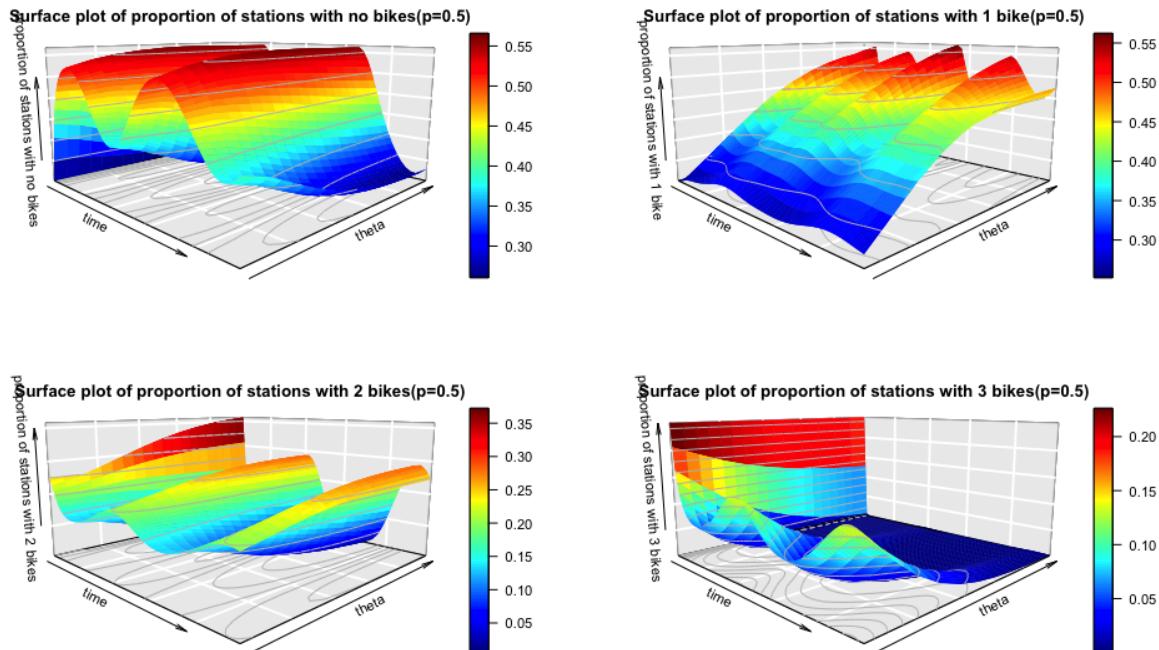

Figure 15: Surface plot of $y_{(t,\theta)}$ on the time-θ plane when $K = 3$ and $p = 0.5$
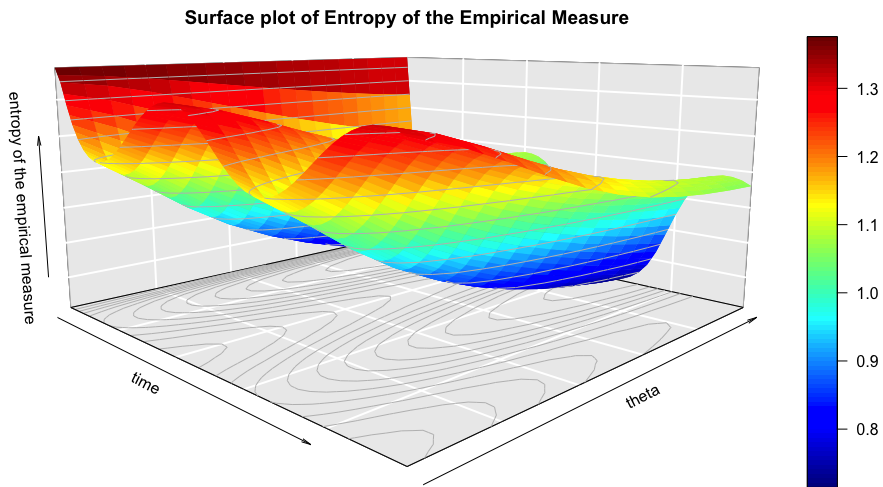
Surface plot of Entropy of the Empirical Measure

Figure 16: Surface plot of entropy of the empirical measure on the time-p plane when $K = 3$

5.2.4 Impact of fleet size (γ) to the empirical measure

In this section, we explore the impact of fleet size on the empirical measure process, as this is very importance from a system design and management point of view.

Let γ be the average number of bikes per station. We use the same parameter setting as in previous section, with $K = 3, p = 0.5$ and exponential choice function with $\theta = 2$. The range of γ is [0.5,2.5]. The initial distribution of bikes is $y(0) = (1 - 0.4\gamma, 0, 0.2\gamma, 0.2\gamma)$.

In Figure 17 we show the surface plot of the fluid limit on the time-γ plane, with $K = 3$, $p = 0.5$ and $\gamma \in [0.5, 2.5]$. We observe that as fleet size (average number of bikes per station) increases, average proportion of empty stations goes down and there is not much change in the variation of proportion of empty stations over time. In terms of full stations, both the average proportion of full stations and the amplitude of its variation over time increase. This indicates that there exists an optimal fleet size γ that minimize proportion of empty and full stations overall.

In Figure 18 we show the surface plot of the entropy of the empirical measure on the time-γ plane. We observe that as fleet size increases, the average entropy increases first then slightly goes down, while the amplitude of the variation of entropy goes down. This indicates that there exists an optimal fleet size γ that minimize the entropy of the empirical measure.

In summary, this example suggests that there exists an optimal value for fleet size that minimize the mean and variance of entropy of the empirical measure under the non-stationary setting. The optimal value can be easily numerically identified with the mean field limit results.
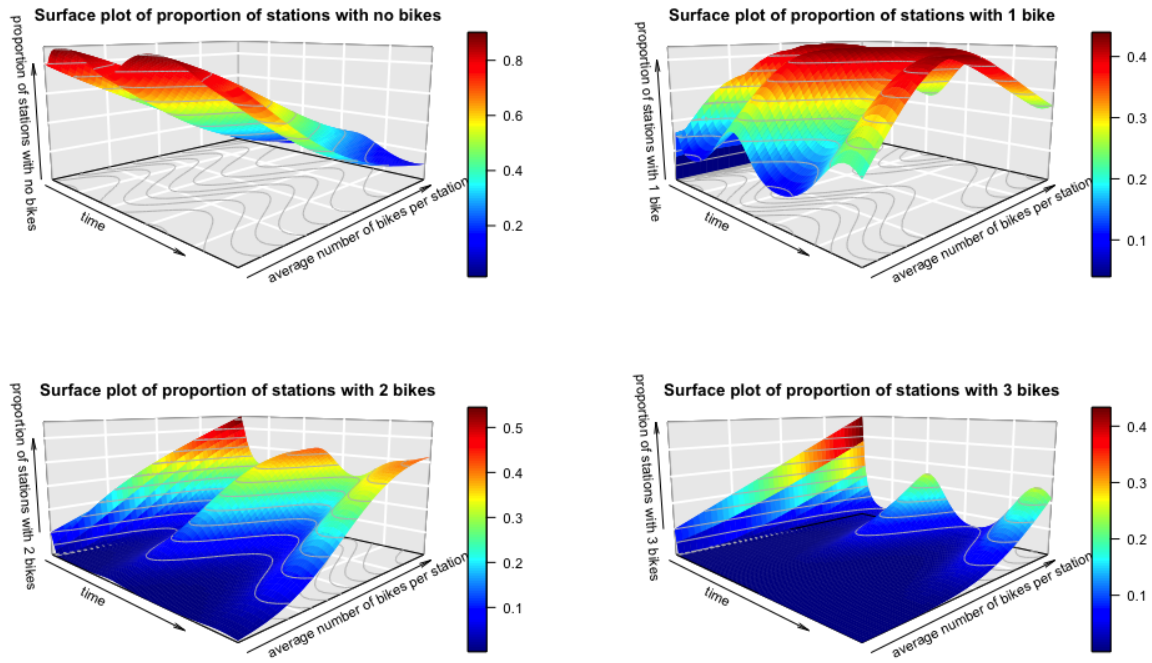
Figure 17: Surface plot of $y_{(t,\gamma)}$ on the time-γ plane when $K = 3, p = 0.5$ and $\theta = 2$
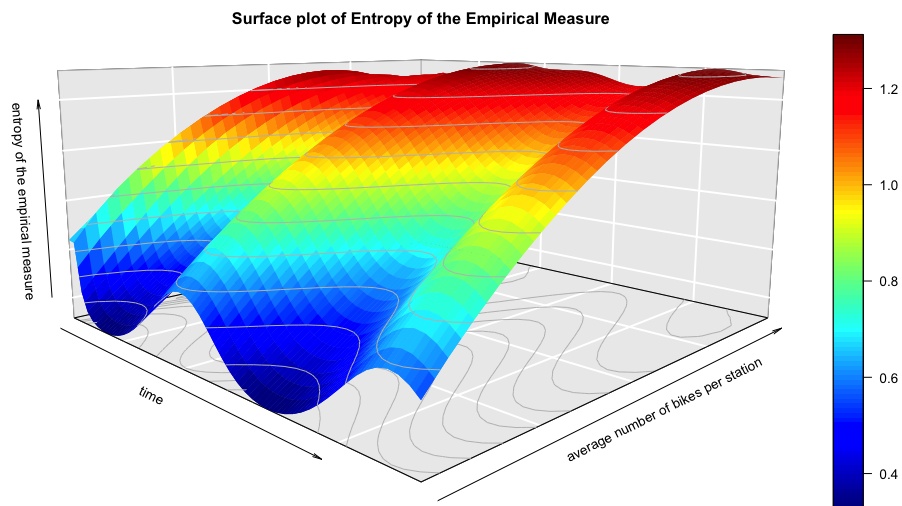

Figure 18: Surface plot of entropy of the empirical measure on the time-p plane when $K = 3$

6 Conclusion

In this paper, we construct a stochastic queueing model with customer choice modeling for the empirical measure process of a finite-capacity bike sharing system. We consider a fraction p of the customers having information to the bike availability of the system, which impacts their arrival rate to the stations through a choice function.

First, we prove a mean field limit and a central limit theorem for an empirical process of the number of stations with k bikes, when station capacities are the same across the network. We also defined a new process called the ratio process in case when station capacities are different, and it characterizes the proportion of stations whose bike availability ratio lies within a particular partition of the interval [0,1]. We prove similar mean field limit and central limit theorem results for the ratio process. These asymptotic results provide great insights into the mean, variance and sample path dynamics of large scale bike sharing systems.

Second, we provide steady state analysis of the mean field limit of both the empirical measure process and the ratio process. We explicitly compute the equilibrium of the mean field limit and prove that interchanging of limits ($\lim_{t\to\infty}\lim_{N\to\infty} = \lim_{N\to\infty}\lim_{t\to\infty}$) holds for the mean field limit in equilibrium.

Finally, we give numerical examples to provide insights on how the fraction of customers with information (p), different choice models, customers' sensitivity to information and fleet size impact the system performance, under both stationary arrivals setting and non-stationary arrivals setting. Our examples show that generally, under stationary arrivals, the more customers with information, the better the system performance is. Under non-stationary settings, there might exists a trade-off between average system performance and its variation over time, and whether more information is more beneficial depends on the specific parameter setting. The form of choice models also plays role in how it impacts the system in terms of controlling empty and full stations.

Despite our analysis, there are many future directions for research.

1. The first direction is showing that the interchanging of limits holds for the diffusion limit as well. This requires proving exponential stability of the mean field limit.

2. We are also interested in constructing a stochastic queueing model for bike sharing systems that takes into account geographic location of stations. In this case we would consider customers arriving via a spatial Poisson process and would consider not only available bikes but also distance to the stations for picking up bikes.

3. Our model can also be adopted to study the effect of customer incentives or reward programs. These programs give riders the opportunity to move bikes from stations to stations, with the hope of improving the overall availability of bikes and docks across the network. For example, some major bike sharing systems such as CitiBike has so-called "Bike Angels", which is a reward program to riders who help move bikes from full stations to empty stations. This may potentially be a more cost-effective way to rebalancing than hiring employees to move bikes. See Chung et al. [3] for more discussion on the topic. In this case, the effect of "Bike Angels" can be modeled similarly to the choice model in this paper as follows,

- Rate of picking up bikes at station i:

$$Q(X_i(t), X_i(t) - 1) = \left(\underbrace{(1-p)\lambda_i}_{\text{no information arrival}} + \underbrace{p(1-\alpha)\lambda_i N \frac{g(X_i(t))}{\sum_{j=1}^{N} g(X_j(t))}}_{\text{with information arrival}} \right.$$

$$\left. + \underbrace{p\alpha\lambda_i N \frac{h(X_i(t))}{\sum_{j=1}^{N} h(X_j(t))}}_{\text{bike angles}} \right) \cdot \underbrace{\mathbf{1}\{X_i(t) > 0\}}_{\text{station } i \text{ not empty}}$$

- Rate of dropping off bikes at station i :

$$Q(X_i(t), X_i(t) + 1) = \underbrace{\mu}_{\text{service rate}} \cdot \left(\underbrace{P_i(1-p\alpha)}_{\text{routing probability}} + \underbrace{P_i p\alpha N \frac{h(-X_i(t))}{\sum_{j=1}^{N} h(-X_j(t))}}_{\text{routing probability for bike angles}} \right) \cdot$$

$$\underbrace{\left(M - \sum_{k=1}^{N} X_k(t) \right)}_{\text{\# of bikes in circulation}} \cdot \underbrace{\mathbf{1}\{X_i(t) < K_i\}}_{\text{station } i \text{ not full}}$$

where g, h are arbitrary functions that are non-decreasing and non-negative.

We intend to pursue these extensions in future work.

7   Appendix

Proof of Theorem 2.3. The proof we give is obtained by Doob's inequality for martingales and Gronwall's lemma. We use Proposition 7.1, Proposition 7.2, and Proposition 7.3 in the proof, and they are stated after Theorem 2.3.

Since Y_t^N is a semi-martingale, we have the following decomposition of Y_t^N,

$$Y_t^N = \underbrace{Y_0^N}_{\text{initial condition}} + \underbrace{M_t^N}_{\text{martingale}} + \int_0^t \underbrace{\beta(Y_s^N)}_{\text{drift term}} ds \qquad (7.55)$$

where Y_0^N is the initial condition and M_t^N is a independent family of martingales. Moreover, $\int_0^t \beta(Y_s^N)ds$ is the integral of the drift term where the drift term is given by $\beta : [0,1]^{K+1} \to$

$\mathbb{R}^{K+1}$ or

$$
\begin{aligned}
\beta(y) \;&=\; \sum_{x \neq y} (x - y) Q(y, x) \\
&=\; \sum_{n=0}^{K} \left[\left((1 - p) + p \frac{g(n)}{\sum_{j=0}^{K} y_j g(j)} \right) (\mathbf{1}_{n-1} - \mathbf{1}_n) \lambda \mathbf{1}_{n>0} \right. \\
&\qquad \left. + \left(\frac{M}{N} - \sum_{j=0}^{K} j y_j \right) (\mathbf{1}_{n+1} - \mathbf{1}_n) \mathbf{1}_{n<K} \right] y_n.
\end{aligned}
\tag{7.56}
$$

We want to compare the empirical measure Y_t^N with the mean field limit y_t defined by

$$
y_t = y_0 + \int_0^t b(y_s)\,ds.
\tag{7.57}
$$

Let $|\cdot|$ denote the Euclidean norm in $\mathbb{R}^{K+1}$, then

$$
\begin{aligned}
\left| Y_t^N - y_t \right| \;&=\; \left| Y_0^N + M_t^N + \int_0^t \beta(Y_s^N)\,ds - y_0 - \int_0^t b(y_s)\,ds \right| \\
&=\; \left| Y_0^N - y_0 + M_s^N + \int_0^t \left(\beta(Y_s^N) - b(Y_s^N) \right) ds + \int_0^t (b(Y_s^N) - b(y_s))\,ds \right|.
\end{aligned}
\tag{7.58}
$$

Now define the random function $f^N(t) = \sup_{s \leq t} \left| Y_s^N - y_s \right|$, we have

$$
f^N(t) \leq |Y_0^N - y_0| + \sup_{s \leq t} |M_s^N| + \int_0^t |\beta(Y_s^N) - b(Y_s^N)|\,ds + \int_0^t |b(Y_s^N) - b(y_s)|\,ds.
$$

By Proposition 7.2, $b(y)$ is Lipschitz with respect to Euclidean norm. Let L be the Lipschitz constant of $b(y)$, then

$$
\begin{aligned}
f^N(t) \;&\leq\; |Y_0^N - y_0| + \sup_{s \leq t} |M_s^N| + \int_0^t |\beta(Y_s^N) - b(Y_s^N)|\,ds + \int_0^t |b(Y_s^N) - b(y_s)|\,ds \\
&\leq\; |Y_0^N - y_0| + \sup_{s \leq t} |M_s^N| + \int_0^t |\beta(Y_s^N) - b(Y_s^N)|\,ds + L \int_0^t |Y_s^N - y_s|\,ds \\
&\leq\; |Y_0^N - y_0| + \sup_{s \leq t} |M_s^N| + \int_0^t |\beta(Y_s^N) - b(Y_s^N)|\,ds + L \int_0^t f^N(s)\,ds.
\end{aligned}
\tag{7.59}
$$

By Gronwall's lemma (See Ames and Pachpatte [1]),

$$
f^N(t) \leq \left(|Y_0^N - y_0| + \sup_{s \leq t} |M_s^N| + \int_0^t |\beta(Y_s^N) - b(Y_s^N)|\,ds \right) e^{Lt}.
\tag{7.60}
$$

Now to bound $f^N(t)$ term by term, we define function $\alpha : [0,1]^{K+1} \to \mathbb{R}^{K+1}$ as

$$
\begin{aligned}
\alpha(y) &= \sum_{x \neq y} |x - y|^2 Q(y, x) \\
&= \frac{1}{N} \sum_n \sum_r \left[\frac{1}{rNR_{\max}} (\mathbf{1}_{(r,n-1)} + \mathbf{1}_{(r,n)}) \mathbf{1}_{n>0} \right. \\
&\quad \left. + \left(\frac{M}{N} - \sum_{n'} \sum_{r'} n' y(r', n') \right) (\mathbf{1}_{(r,n+1)} + \mathbf{1}_{(r,n)}) \mathbf{1}_{n<K} \right] \cdot y(r, n) \qquad (7.61)
\end{aligned}
$$

and consider the following four sets

$$
\begin{aligned}
\Omega_0 &= \{|Y_0^N - y_0| \leq \delta\}, &(7.62) \\
\Omega_1 &= \left\{ \int_0^{t_0} |\beta(Y_s^N) - b(Y_s^N)| ds \leq \delta \right\}, &(7.63) \\
\Omega_2 &= \left\{ \int_0^{t_0} \alpha(Y_t^N) dt \leq A(N) t_0 \right\}, &(7.64) \\
\Omega_3 &= \left\{ \sup_{t \leq t_0} |M_t^N| \leq \delta \right\}, &(7.65)
\end{aligned}
$$

where $\delta = \epsilon e^{-Lt_0}/3$. Here the set Ω_1 is to bound the initial condition, the set Ω_2 is to bound the drift term β and the limit of drift term b, and the sets Ω_2, Ω_3 are to bound the martingale M_t^N. Therefore on the event $\Omega_0 \cap \Omega_1 \cap \Omega_3$,

$$
f^N(t_0) \leq 3\delta e^{Lt_0} = \epsilon. \qquad (7.66)
$$

Since $\lim_{N \to \infty} \frac{M}{N} = \gamma$ and $\lim_{N \to \infty} NR_{\max} = \frac{1}{\Lambda}$, we can choose large enough N such that

$$
\frac{M}{N} \leq 2\gamma, \quad NR_{\max} \geq \frac{1}{2\Lambda}.
$$

Thus

$$
\begin{aligned}
\alpha(y) &= \sum_{x \neq y} |x - y|^2 Q(y, x) \\
&= \frac{1}{N} \sum_{n=0}^K \left[\left((1-p) + p \frac{g(n)}{\sum_{j=0}^K y_j g(j)} \right) (\mathbf{1}_{n-1} + \mathbf{1}_n) \lambda \mathbf{1}_{n>0} \right. \\
&\quad \left. + \left(\frac{M}{N} - \sum_{j=0}^K j y_j \right) (\mathbf{1}_{n+1} + \mathbf{1}_n) \mathbf{1}_{n<K} \right] y_n \\
&\leq \frac{1}{N} \sum_{n=0}^K \left[\left((1-p) + p \frac{C_{\max}}{C_{\min}} \right) \cdot 2\lambda + 2\gamma \cdot 2 \right] y_n \\
&= \frac{2}{N} \left[\left((1-p) + p \frac{C_{\max}}{C_{\min}} \right) \lambda + 2\gamma \right] \\
&\sim O(\frac{1}{N}). \qquad (7.67)
\end{aligned}
$$

Consider the stopping time

$$T = t_0 \wedge \inf \left\{ t \geq 0 : \int_0^t \alpha(Y_s^N) ds > A(N)t_0 \right\},$$

by Proposition 7.1,

$$\mathbb{E} \left(\sup_{t \leq T} |M_t^N|^2 \right) \leq 4\mathbb{E} \int_0^T \alpha(Y_t^N) dt \leq 4A(N)t_0.$$

On Ω_2, we have $T = t_0$, so $\Omega_2 \cap \Omega_3^c \subset \{\sup_{t \leq T} |M_t^N| > \delta\}$. By Chebyshev's inequality,

$$\mathbb{P}(\Omega_2 \cap \Omega_3^c) \leq \mathbb{P} \left(\sup_{t \leq T} |M_t^N| > \delta \right) \leq \frac{\mathbb{E} \left(\sup_{t \leq T} |M_t^N|^2 \right)}{\delta^2} \leq 4A(N)t_0/\delta^2. \tag{7.68}$$

Thus by Equation (7.66), we have the following result,

$$\begin{aligned}
\mathbb{P} \left(\sup_{t \leq t_0} |Y_t^N - y_t| > \epsilon \right) &\leq \mathbb{P}(\Omega_0^c \cup \Omega_1^c \cup \Omega_3^c) \\
&\leq \mathbb{P}(\Omega_2 \cap \Omega_3^c) + \mathbb{P}(\Omega_0^c \cup \Omega_1^c \cup \Omega_2^c) \\
&\leq 4A(N)t_0/\delta^2 + \mathbb{P}(\Omega_0^c \cup \Omega_1^c \cup \Omega_2^c) \\
&= 36A(N)t_0 e^{2Lt_0}/\epsilon^2 + \mathbb{P}(\Omega_0^c \cup \Omega_1^c \cup \Omega_2^c).
\end{aligned} \tag{7.69}$$

Let $A(N) = \frac{4(C+\gamma)}{N}$, then $\Omega_2^c = \emptyset$. And since $Y_0^N \xrightarrow{p} y_0$, $\lim_{N \to \infty} \mathbb{P}(\Omega_2^c) = 0$. Therefore we have

$$\lim_{N \to \infty} \mathbb{P} \left(\sup_{t \leq t_0} |Y_t^N - y_t| > \epsilon \right) = \lim_{N \to \infty} \mathbb{P}(\Omega_1^c).$$

By Proposition 7.3, $\lim_{N \to \infty} \mathbb{P}(\Omega_1^c) = 0$. Thus, we proved the final result

$$\lim_{N \to \infty} \mathbb{P} \left(\sup_{t \leq t_0} |Y_t^N - y_t| > \epsilon \right) = 0.$$

$\square$

Proposition 7.1 (Bounding martingales)**.** *For any stopping time T such that $\mathbb{E}(T) < \infty$, we have*

$$\mathbb{E} \left(\sup_{t \leq T} |M_t^N|^2 \right) \leq 4\mathbb{E} \int_0^T \alpha(Y_t^N) dt. \tag{7.70}$$

Proof. Let $\tilde{\mu}$ be the jump measure of Y_t^N, and ν be its compensator, defined on $(0, \infty) \times [0, 1]$ by

$$\tilde{\mu} = \sum_{t : Y_t^N \neq Y_{t-}^N} \delta(t, Y_t^N), \qquad \nu(dt, B) = Q(Y_{t-}^N, B)dt \quad \forall B \in \mathcal{B}([0, 1]). \tag{7.71}$$

Let $\tilde{Y}^N_{m,t}$ be the jump chain of Y^N_t, J_m be the jump time, then we have for any $t \in [0, \infty)$, $J_n \leq t < J_{n+1}$ for some $n \geq 0$. The martingale M^N_t can be written as

$$
\begin{aligned}
M^N_t &= Y^N_t - Y^N_0 - \int_0^t \beta(Y^N_s) ds \\
&= \sum_{m=0}^{n-1} (\tilde{Y}^N_{m+1} - \tilde{Y}^N_m) - \int_0^t \int_0^1 (y - Y^N_{s-}) Q(Y^N_{s-}, dy) ds \\
&= \int_0^t \int_0^1 (y - Y^N_{s-}) \tilde{\mu}(ds, dy) - \int_0^t \int_0^1 (y - Y^N_{s-}) \nu(ds, dy) \\
&= \int_0^t \int_0^1 (y - Y^N_{s-})(\tilde{\mu} - \nu)(ds, dy). \quad (7.72)
\end{aligned}
$$

Note the following identity

$$
(M^N_t)^2 = 2 \int_0^t \int_0^1 M_{s-}(y - Y^N_{s-})(\tilde{\mu} - \nu)(ds, dy) + \int_0^t \int_0^1 (y - Y^N_{s-})^2 \tilde{\mu}(ds, dy). \quad (7.73)
$$

This can be established by verifying that the jumps of the left and right hand sides agree, and that their derivatives agree between jump times. Then we can write

$$
(M^N_t)^2 = N^N_t + \int_0^t \alpha(Y^N_t) ds \quad (7.74)
$$

where

$$
N^N_t = \int_0^t \int_0^1 H(s, y)(\tilde{\mu} - \nu)(ds, dy), \quad (7.75)
$$

and

$$
H(s, y) = 2 M_{s-}(y - Y^N_{s-}) + (y - Y^N_{s-})^2. \quad (7.76)
$$

Consider the previsible process

$$
H_2(t, y) = H(t, y) \mathbf{1}_{\{t \leq T \wedge T_n\}} \quad (7.77)
$$

where $T_n = \inf\{t \geq 0 : \beta(Y^N_t) > n\} \wedge n$. Then

$$
N_{T \wedge T_n} = \int_0^\infty \int_0^1 H_2(t, y)(\tilde{\mu} - \nu)(dt, dy), \quad (7.78)
$$

and

$$
\begin{aligned}
\mathbb{E} \int_0^\infty \int_0^1 |H_2(s, y)| \nu(ds, dy) &= \mathbb{E} \int_0^{T \wedge T_n} \int_0^1 |2 M_{s-}(y - Y^N_{s-}) + (y - Y^N_{s-})^2| \nu(ds, dy) \\
&= \mathbb{E} \int_0^{T \wedge T_n} \left(2|M^N_t| \beta(Y^N_t) + \alpha(Y^N_t) \right) dt \\
&\leq \mathbb{E} \int_0^{T \wedge T_n} 2(2 + n^2) n \, dt + \mathbb{E} \int_0^T \alpha(Y^N_t) dt \\
&\leq 2n^4 + 4n^2 + \frac{4(C + \gamma)}{N} \mathbb{E}(T) < \infty. \quad (7.79)
\end{aligned}
$$

By Theorem 8.4 in Darling et al. [5], we have that $N^{T \wedge T_n}$ is a martingale. Replace t by $T \wedge T_n$ in Equation (7.74) and take expectation to obtain

$$\mathbb{E}(|M_{T \wedge T_n}|^2) = \mathbb{E}(N_{T \wedge T_n}) + \mathbb{E} \int_0^{T \wedge T_n} \alpha(Y_t^N)dt. \qquad (7.80)$$

Since $N^{T \wedge T_n}$ is a martingale,

$$\mathbb{E}(N_{T \wedge T_n}) = \mathbb{E}(N_0) = 0. \qquad (7.81)$$

Apply Doob's L^2-inequality to the martingale $M^{T \wedge T_n}$ to obtain

$$\mathbb{E}\left(\sup_{t \le T \wedge T_n} |M_t|^2\right) \le 4\mathbb{E}(|M_{T \wedge T_n}|^2)$$

$$= 4\mathbb{E} \int_0^{T \wedge T_n} \alpha(Y_t^N)dt. \qquad (7.82)$$

$\square$

Proposition 7.2 (Asymptotic Drift is Lipschitz). *The drift function $b(y)$ given in Equation (2.7) is a Lipschitz function with respect to the Euclidean norm in $\mathbb{R}^{K+1}$.*

Proof. Denote $|\cdot|$ the Euclidean norm in $\mathbb{R}^{K+1}$. Consider $y, \tilde{y} \in [0,1]^{K+1}$,

$$|b(y) - b(\tilde{y})| \le 2\lambda(1-p)|y - \tilde{y}| + \lambda p \left(\sum_{k=0}^K \left|\frac{y_k g(k)}{\sum_{j=0}^K y_j g(j)} - \frac{\tilde{y}_k g(k)}{\sum_{j=0}^K \tilde{y}_j g(j)}\right|^2\right)^{1/2} + 2\gamma|y - \tilde{y}|$$

$$\le 2(\lambda(1-p) + \gamma)|y - \tilde{y}| + \lambda p \left(\sum_{k=0}^K \left(\frac{g(k)(y_k - \tilde{y}_k)}{\min_{0 \le j \le K} g(j)}\right)^2\right)^{1/2}. \qquad (7.83)$$

Since $\{g_j\}$ are non-negative, there exists constants $C_{\min}, C_{\max} > 0$ such that

$$C_{\min} \le \min_{0 \le j \le K} g(j) \le \max_{0 \le j \le K} g(j) \le C_{\max}. \qquad (7.84)$$

Then

$$|b(y) - b(\tilde{y})| \le 2(\lambda(1-p) + \gamma)|y - \tilde{y}| + \lambda p \frac{C_{\max}}{C_{\min}} \left(\sum_{k=0}^K (y_k - \tilde{y}_k)^2\right)^{1/2}$$

$$= 2\left(\lambda(1-p) + \lambda p \frac{C_{\max}}{C_{\min}} + \gamma\right)|y - \tilde{y}| \qquad (7.85)$$

which proves that $b(y)$ is Lipschitz with respect to Euclidean norm in $\mathbb{R}^{K+1}$. $\square$

Proposition 7.3 (Drift is Asymptotically Close to Lipschitz Drift). *Under the assumptions of Theorem 2.3, we have for any $\epsilon > 0$ and $s \ge 0$,*

$$\lim_{N \to \infty} P(|\beta(Y_s^N) - b(Y_s^N)| > \epsilon) = 0.$$

Proof.

$$\begin{aligned}
\left|\beta(Y_s^N) - b(Y_s^N)\right| &= \left|\sum_{n=0}^{K}\left(\frac{M}{N} - \gamma\right)(\mathbf{1}_{n+1} - \mathbf{1}_n)\mathbf{1}_{n<K}Y_s^N(n)\right| \\
&\leq 2\left|\frac{M}{N} - \gamma\right|\left|\sum_{n=0}^{K}Y_s^N(n)\right| \\
&= 2\left|\frac{M}{N} - \gamma\right| \to 0. \tag{7.86}
\end{aligned}$$

$\square$

Lemma 7.4. $\sqrt{N}M_t^N$ *is a family of martingales independent of D_0^N with Doob-Meyer brackets given by*

$$\left\langle \sqrt{N}M^N(k)\right\rangle_t = \int_0^t (\beta_+(Y_s^N)(k) + \beta_-(Y_s^N)(k))ds,$$

$$\left\langle \sqrt{N}M^N(k), \sqrt{N}M^N(k+1)\right\rangle_t = -\int_0^t\left[\left((1-p) + p\frac{g(k+1)}{\sum_{j=0}^K jy_j}\right)\lambda Y_s^N(k+1)\right.$$
$$\left. + \left(\frac{M}{N} - \sum_{j=0}^K jY_s^N(j)\right)Y_s^N(k)\right]ds \quad \text{for } k < K,$$

$$\left\langle \sqrt{N}M^N(k), \sqrt{N}M^N(j)\right\rangle_t = 0 \quad \text{for } |k-j| > 1. \tag{7.87}$$

Proof. By Dynkin's formula,

$$\begin{aligned}
\left\langle \sqrt{N}M^N(k)\right\rangle_t &= \int_0^t N\sum_{x\neq Y_s^N}|x(k) - Y_s^N(k)|^2 Q(Y_s^N, x)ds \\
&= N\int_0^t \alpha(Y_s^N)(k)ds \\
&= \int_0^t \sum_r \left[\frac{1}{rNR_{\max}}\left(Y_s^N(r,k+1)\mathbf{1}_{k<K} + Y_s^N(r,k)\mathbf{1}_{k>0}\right)\right. \\
&\quad + \left(\frac{M}{N} - \sum_{n'}\sum_{r'}n'Y_s^N(r',n')\right)\left(Y_s^N(r,k)\mathbf{1}_{k<K} + Y_s^N(r,k-1)\mathbf{1}_{k>0}\right)\right]ds \\
&= \int_0^t (\beta_+(Y_s^N)(k) + \beta_-(Y_s^N)(k))ds.
\end{aligned}$$

$$\tag{7.88}$$

To compute $\left\langle \sqrt{N}M^N(k), \sqrt{N}M^N(k+1)\right\rangle_t$ for $k < K$, since

$$\left\langle M^N(k) + M^N(k+1)\right\rangle_t$$

$$= \int_0^t \sum_{x \neq Y_s^N} \left|x(k) + x(k+1) - Y_s^N(k) - Y_s^N(k+1)\right|^2 Q(Y_s^N, x)ds$$

$$= \frac{1}{N}\int_0^t \sum_r \left[\frac{1}{rNR_{\max}}\left(Y_s^N(r, k+2)\mathbf{1}_{k<K-1} + Y_s^N(r,k)\mathbf{1}_{k>0}\right)\right.$$

$$\left. + \left(\frac{M}{N} - \sum_{n'}\sum_{r'}n'Y_s^N(r',n')\right)\left(Y_s^N(r,k+1)\mathbf{1}_{k<K-1} + Y_s^N(r,k-1)\mathbf{1}_{k>0}\right)\right]ds. \tag{7.89}$$

We have that

$$\left\langle \sqrt{N}M^N(k), \sqrt{N}M^N(k+1)\right\rangle_t$$

$$= \frac{N}{2}\left[\left\langle M^N(k) + M^N(k+1)\right\rangle_t - \left\langle M^N(k)\right\rangle_t - \left\langle M^N(k+1)\right\rangle_t\right]$$

$$= \frac{1}{2}\int_0^t \sum_r \left[\frac{1}{rNR_{\max}}\left(Y_s^N(r,k+2)\mathbf{1}_{k<K-1} + Y_s^N(r,k)\mathbf{1}_{k>0}\right)\right.$$

$$\left. + \left(\frac{M}{N} - \sum_{n'}\sum_{r'}n'Y_s^N(r',n')\right)\left(Y_s^N(r,k+1)\mathbf{1}_{k<K-1} + Y_s^N(r,k-1)\mathbf{1}_{k>0}\right)\right]ds \tag{7.90}$$

$$- \frac{1}{2}\int_0^t (\beta_+(Y_s^N)(k) + \beta_+(Y_s^N)(k+1) + \beta_-(Y_s^N)(k) + \beta_-(Y_s^N)(k+1))ds$$

$$= -\int_0^t \sum_r \left[\frac{1}{rNR_{\max}}Y_s^N(r,k+1) + \left(\frac{M}{N} - \sum_{n'}\sum_{r'}n'Y_s^N(r',n')\right)Y_s^N(r,k)\right]ds.$$

When $|k - j| > 1$, $M^N(k)$ and $M^N(j)$ are independent, thus

$$\left\langle \sqrt{N}M^N(k), \sqrt{N}M^N(j)\right\rangle_t = 0. \tag{7.91}$$

$\square$

Proposition 7.5. *For any $s \geq 0$,*

$$\limsup_{N \to \infty} \sqrt{N}\left|\beta(Y_s^N) - b(Y_s^N)\right| < \infty. \tag{7.92}$$

Proof.

$$\limsup_{N \to \infty} \sqrt{N}\left|\beta(Y_s^N) - b(Y_s^N)\right| = \limsup_{N \to \infty} \sqrt{N}\left|\sum_{n=0}^K \left(\frac{M}{N} - \gamma\right)(\mathbf{1}_{n+1} - \mathbf{1}_n)\mathbf{1}_{n<K}Y_s^N(n)\right|$$

$$\leq \limsup_{N \to \infty} 2\sqrt{N}\left|\frac{M}{N} - \gamma\right|\left|\sum_{n=0}^K Y_s^N(n)\right|$$

$$= \limsup_{N \to \infty} 2\sqrt{N}\left|\frac{M}{N} - \gamma\right|$$

$$< \infty. \tag{7.93}$$

□

Lemma 7.6. *(Finite Horizon Bound) For any $T \geq 0$, if*

$$\limsup_{N \to \infty} \mathbb{E}\left(|D_0^N|^2\right) < \infty,$$

then we have

$$\limsup_{N \to \infty} \mathbb{E}\left(\sup_{0 \leq t \leq T} |D_t^N|^2\right) < \infty.$$

Proof. By Proposition 7.5, $\sqrt{N}|\beta(Y_s^N) - b(Y_s^N)| = O(1)$, then

$$\begin{aligned}
|D_t^N| &\leq |D_0^N| + \sqrt{N}|M_t^N| + O(1)t + \int_0^t \sqrt{N}|b(Y_s^N) - b(y_s)|ds \\
&\leq |D_0^N| + \sqrt{N}|M_t^N| + O(1)t + \int_0^t \sqrt{N}L|Y_s^N - y_s|ds \qquad (7.94) \\
&= |D_0^N| + \sqrt{N}|M_t^N| + O(1)t + \int_0^t L|D_s^N|ds.
\end{aligned}$$

By Gronwall's Lemma,

$$\sup_{0 \leq t \leq T} |D_t^N| \leq e^{LT}\left(|D_0^N| + O(1)T + \sup_{0 \leq t \leq T}|\sqrt{N}M_t^N|\right),$$

then

$$\limsup_{N \to \infty} \mathbb{E}\left(\sup_{0 \leq t \leq T} |D_t^N|^2\right) \leq e^{2LT}\left[\limsup_{N \to \infty} \mathbb{E}(|D_0^N|) + O(1)T + \limsup_{N \to \infty} \mathbb{E}\left(\sup_{0 \leq t \leq T} \sqrt{N}|M_t^N|\right)\right]^2.$$

We know that

$$\left[\mathbb{E}\left(\sup_{0 \leq t \leq T} \sqrt{N}|M_t^N|\right)\right]^2 \leq N\mathbb{E}\left(\sup_{0 \leq t \leq T}|M_t^N|^2\right) \leq 4NA(N)T,$$

and that $A(N) = O(\frac{1}{N})$. Therefore

$$\limsup_{N \to \infty} \mathbb{E}\left(\sup_{0 \leq t \leq T} \sqrt{N}|M_t^N|\right) < \infty.$$

Together with our assumption $\limsup_{N \to \infty} \mathbb{E}(|D_0^N|^2) < \infty$, we have

$$\limsup_{N \to \infty} \mathbb{E}\left(\sup_{0 \leq t \leq T} |D_t^N|^2\right) < \infty.$$

□

Lemma 7.7. *If $(D_0^N)_{N=1}^\infty$ is tight then $(D^N)_{N=1}^\infty$ is tight and its limit points are continuous.*

Proof. To prove the tightness of $(D^N)_{N=1}^\infty$ and the continuity of the limit points, we only need to show the following two conditions holds for each $T > 0$ and $\epsilon > 0$,

(i)

$$\lim_{K \to \infty} \limsup_{N \to \infty} \mathbb{P}\left(\sup_{0 \leq t \leq T} |D_t^N| > K \right) = 0, \qquad (7.95)$$

(ii)

$$\lim_{\delta \to 0} \limsup_{N \to \infty} \mathbb{P}\left(w(D^N, \delta, T) \geq \epsilon \right) = 0 \qquad (7.96)$$

where for $x \in \mathbb{D}^d$,

$$w(x, \delta, T) = \sup\left\{ \sup_{u,v \in [t,t+\delta]} |x(u) - x(v)| : 0 \leq t \leq t + \delta \leq T \right\}. \qquad (7.97)$$

By Lemma 7.6, there exists $C_0 > 0$ such that

$$\begin{aligned}
\lim_{K \to \infty} \limsup_{N \to \infty} \mathbb{P}\left(\sup_{0 \leq t \leq T} |D_t^N| > K \right) &\leq \lim_{K \to \infty} \limsup_{N \to \infty} \frac{\mathbb{E}\left(\sup_{0 \leq t \leq T} |D_t^N|^2 \right)}{K^2} \\
&\leq \lim_{K \to \infty} \frac{C_0}{K^2} \\
&= 0, \qquad (7.98)
\end{aligned}$$

which proves condition (i).

For condition (ii), we have that

$$\begin{aligned}
D_u^N - D_v^N &= \underbrace{\sqrt{N} \cdot (M_u^N - M_v^N)}_{\text{first term}} + \underbrace{\int_v^u \sqrt{N} \left(\beta(Y_z^N) - b(Y_z^N) \right) dz}_{\text{second term}} \\
&+ \underbrace{\int_v^u \sqrt{N} \left(b(Y_z^N) - b(y_z) \right) dz}_{\text{third term}} \qquad (7.99)
\end{aligned}$$

for any $0 < t \leq u < v \leq t + \delta \leq T$. Now it suffices to show that each of the three terms of $D_u^N - D_v^N$ satisfies condition (ii). In what follows, we will show that each of the three terms satisfies condition (ii) to complete the proof of tightness.

For the first term, similar to the proof of Proposition 7.3, we can show that

$$\sup_{t \leq T} \left| \beta_+(Y_t^N) - b_+(Y_t^N) \right| \xrightarrow{p} 0, \quad \sup_{t \leq T} \left| \beta_-(Y_t^N) - b_-(Y_t^N) \right| \xrightarrow{p} 0.$$

And by the proof of Proposition 7.2, $b_+(y), b_-(y)$ are also Lipschitz with constant L, then by the fact that the composition of Lipschitz functions are also Lipschitz,

$$\max\left\{ \sup_{t \leq T} |b_+(Y_t^N) - b_+(y_t)|, \sup_{t \leq T} |b_-(Y_t^N) - b_-(y_t)| \right\} \leq L \sup_{t \leq T} |Y_t^N - y_t|. \qquad (7.100)$$

47

By Theorem 2.3,

$$\sup_{t \leq T} |Y_t^N - y_t| \xrightarrow{p} 0. \tag{7.101}$$

Thus, for any $\epsilon > 0$,

$$
\begin{aligned}
&\lim_{N \to \infty} \mathbb{P} \left(\sup_{t \leq T} \left| \left\langle \sqrt{N} M^N(k) \right\rangle_t - \langle M(k) \rangle_t \right| > \epsilon \right) \\
&= \lim_{N \to \infty} \mathbb{P} \left(\sup_{t \leq T} \left| \int_0^t \left(\beta_+(Y_s^N) + \beta_-(Y_s^N) - b_+(y_s) - b_-(y_s) \right) ds \right| > \epsilon \right) \\
&\leq \lim_{N \to \infty} \mathbb{P} \left(\sup_{t \leq T} T \left| \beta_+(Y_t^N) - b_+(Y_t^N) \right| > \epsilon/3 \right) + \lim_{N \to \infty} \mathbb{P} \left(\sup_{t \leq T} T \left| \beta_-(Y_t^N) - b_-(Y_t^N) \right| > \epsilon/3 \right) \\
&\quad + \lim_{N \to \infty} \mathbb{P} \left(\sup_{t \leq T} 2LT \left| Y_t^N - y_t \right| > \epsilon/3 \right) \\
&= 0, \tag{7.102}
\end{aligned}
$$

which implies

$$\sup_{t \leq T} \left| \left\langle \sqrt{N} M^N(k) \right\rangle_t - \langle M(k) \rangle_t \right| \xrightarrow{p} 0. \tag{7.103}$$

We also know that the jump size of D_t^N is $1/\sqrt{N}$, therefore

$$\lim_{N \to \infty} \mathbb{E} \left[\sup_{0 < t \leq T} \left| M_t^N - M_{t-}^N \right| \right] = 0. \tag{7.104}$$

By Theorem 1.4 in Chapter 7 of Ethier and Kurtz [7], $\sqrt{N} M_t^N$ converges to the Brownian motion M_t in distribution in $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^{K+1})$. By Prohorov's theorem, $(\sqrt{N} M^N)_{N=1}^\infty$ is tight. And since M_t is a Brownian motion, its sample path is almost surely continuous.

For the second term, we have by Proposition 7.5 that the quantity $\sqrt{N} \left(\beta(Y_z^N) - b(Y_z^N) \right)$ is bounded for any value of $z \in [0, T]$. Therefore, there exists some constant C_1 that does not depend on N such that

$$\sup_{z \in [0,T]} \sqrt{N} \left| \beta(Y_z^N) - b(Y_z^N) \right| \leq C_1. \tag{7.105}$$

Then

$$
\begin{aligned}
&\lim_{\delta \to 0} \lim_{N \to \infty} \mathbb{P} \left(\sup_{u,v \in [0,T], |u-v| \leq \delta} \int_v^u \sqrt{N} \left| \beta(Y_z^N) - b(Y_z^N) \right| dz > \epsilon \right) \\
&\leq \lim_{\delta \to 0} \lim_{N \to \infty} \mathbb{P} \left(\delta \sup_{z \in [0,T]} \sqrt{N} \left| \beta(Y_z^N) - b(Y_z^N) \right| > \epsilon \right) \\
&\leq \lim_{\delta \to 0} \mathbb{P} \left(\delta C_1 > \epsilon \right) \\
&= 0. \tag{7.106}
\end{aligned}
$$

Thus, we have proved the oscillation bound for the second term. Finally for the third term we have that

$$
\begin{aligned}
\int_v^u \sqrt{N}\left|b(Y_z^N) - b(y_z)\right| dz \;&\leq\; \int_v^u \sqrt{N} L \left|Y_z^N - y_z\right| dz \\
&=\; \int_v^u L \cdot \left|D_z^N\right| dz \\
&\leq\; L\delta \sup_{t\in[0,T]} |D_t^N|.
\end{aligned} \tag{7.107}
$$

By Lemma 7.6,

$$
\begin{aligned}
\lim_{\delta\to 0}\lim_{N\to\infty} \;&\mathbb{P}\left(\sup_{u,v\in[0,T],|u-v|\leq\delta} \int_v^u \sqrt{N}\left|b(Y_z^N) - b(y_z)\right| dz > \epsilon\right) \\
\leq\; &\lim_{\delta\to 0}\lim_{N\to\infty} \mathbb{P}\left(L\delta \sup_{t\in[0,T]} |D_t^N| > \epsilon\right) \\
\leq\; &\lim_{\delta\to 0}\lim_{N\to\infty} \frac{\mathbb{E}\left(\sup_{t\in[0,T]} |D_t^N|^2\right)}{(\epsilon/L\delta)^2} \\
\leq\; &\lim_{\delta\to 0} \frac{C_0 (L\delta)^2}{\epsilon^2} \\
=\; &0,
\end{aligned} \tag{7.108}
$$

which implies that the oscillation bound holds for the third term. $\qquad\square$

Proof of Theorem 4.2. To show that g is a Lyapunov function for the dynamical system $y(t)$, we only need to show that $\frac{d}{dt}g(y(t)) = b(y)\nabla g(y) \leq 0$ and $\frac{d}{dt}g(y(t)) = 0$ if and only if $y = \bar{y}$.
Since

$$
\begin{aligned}
g(y) \;&=\; \sum_{n=0}^{K} y_n \log\left(\frac{y_n}{\nu_{\rho(y)}(n)}\right) - \log(Z(\rho(y))) + \sum_{n=0}^{K}\sum_{k=0}^{n-1} S_{k,n}(y_n) \\
&=\; \sum_{n=0}^{K} y_n \log\left(\frac{y_n Z(\rho(y))}{\prod_{i=0}^{n-1} \rho_i(y)}\right) - \log(Z(\rho(y))) + \sum_{n=0}^{K}\sum_{k=0}^{n-1} S_{k,n}(y_n) \\
&=\; \sum_{n=0}^{K} y_n \log\left(\frac{y_n}{\prod_{i=0}^{n-1} \rho_i(y)}\right) + \sum_{n=0}^{K} y_n \log(Z(\rho(y)) - \log(Z(\rho(y)) + \sum_{n=0}^{K}\sum_{k=0}^{n-1} S_{k,n}(y_n) \\
&=\; \sum_{n=0}^{K} y_n \left(\log(y_n) - \log\left(\prod_{i=0}^{n-1} \rho_i(y)\right)\right) + \sum_{n=0}^{K}\sum_{k=0}^{n-1} S_{k,n}(y_n).
\end{aligned} \tag{7.109}
$$

49

Then we have the gradient of g is

$$
\begin{aligned}
\frac{\partial g(y)}{\partial y_n} &= \log(y_n) + 1 - \sum_{n=0}^{K} y_n \left(\sum_{i=0}^{n-1} \frac{\partial \rho_i(y)}{\partial y_n} / \rho_i(y) \right) - \log \left(\prod_{i=0}^{n-1} \rho_i(y) \right) \\
&\quad + \sum_{n=0}^{K} \sum_{k=0}^{n-1} y_n \phi'_{k,n}(y_n) / \phi_{k,n}(y_n) \\
&= \log \left(\frac{y_n}{\prod_{i=0}^{n-1} \rho_i(y)} \right) + 1.
\end{aligned}
\tag{7.110}
$$

Now denote $b(y) = y B_y$ where B_y is the $(K+1) \times (K+1)$ generator matrix that depends on y. Due to the reversibility of generator B_y with respect to distribution $\nu_{\rho(y)}$, we have

$$
\nu_{\rho(y)}(m) B_y(m, n) = \nu_{\rho(y)}(n) B_y(n, m).
$$

Denote $q_y(m, n) = \nu_{\rho(y)}(m) B_y(m, n)$, then $q_y(m, n)$ is non-negative and symmetric in (m, n). Then we can show that the Dirichlet form of associated to B_y evaluated at vectors $\left(\frac{y_n}{\nu_{\rho(y)}(n)} \right)$ and any $u \in \mathbb{R}^{K+1}$ is equal to $-y B_y u$,

$$
\begin{aligned}
&\mathcal{E} \left(\left(\frac{y_n}{\nu_{\rho(y)}(n)} \right)_n, u \right) \\
&= \frac{1}{2} \sum_{m,n=0}^{K} q_y(m, n) \left(\frac{y_m}{\nu_{\rho(y)}(m)} - \frac{y_n}{\nu_{\rho(y)}(n)} \right) (u_m - u_n) \\
&= \frac{1}{2} \sum_{m,n=0}^{K} \nu_{\rho(y)}(m) B_y(m, n) \left(\frac{y_m u_m}{\nu_{\rho(y)}(m)} + \frac{y_n u_n}{\nu_{\rho(y)}(n)} - \frac{y_m u_n}{\nu_{\rho(y)}(m)} - \frac{y_n u_m}{\nu_{\rho(y)}(n)} \right) \\
&= \frac{1}{2} \sum_{m,n=0}^{K} (B(m, n) y_m u_m + B(m, n) y_m u_m - B(m, n) y_m u_n - B(n, m) y_n u_m) \\
&= \sum_{m=0}^{K} \left(\sum_{n=0}^{K} B(m, n) \right) y_m u_m - \sum_{m,n=0}^{K} B(m, n) y_m u_n \\
&= - \sum_{m,n=0}^{K} B(m, n) y_m u_n \\
&= -y B_y u.
\end{aligned}
\tag{7.111}
$$

Here we use the fact that row sum of B_y is zero. Then using Equation (7.110), we have

$$
y B_y \nabla g(y) = -\frac{1}{2} \sum_{m,n=0}^{K} q_y(m, n) \left(\frac{y_m}{\nu_{\rho(y)}(m)} - \frac{y_n}{\nu_{\rho(y)}(n)} \right) \left(\log \left(\frac{y_m}{\nu_{\rho(y)}(m)} \right) - \log \left(\frac{y_n}{\nu_{\rho(y)}(n)} \right) \right).
\tag{7.112}
$$

This shows that $\frac{d}{dt} g(y(t)) = y B_y \nabla g(y) \le 0$ for all y. It is zero only if $\frac{y_m}{\nu_{\rho(y)}(m)} = \frac{y_n}{\nu_{\rho(y)}(n)}$ for all m, n pairs that $q_y(m, n) > 0$. By the irreducibility of B_y, this is only possible if $\frac{y_n}{\nu_{\rho(y)}(n)}$ doesn't depend on n. This indicates $y = \nu_{\rho(y)}$, which is the equilibrium point of $y(t)$. $\quad \square$

References

[1] William F Ames and BG Pachpatte. *Inequalities for differential and integral equations*, volume 197. Academic press, 1997.

[2] Daniel Chemla, Frédéric Meunier, Thomas Pradeau, Roberto Wolfler Calvo, and Houssame Yahiaoui. Self-service bike sharing systems: simulation, repositioning, pricing. 2013.

[3] Hangil Chung, Daniel Freund, and David B. Shmoys. Bike angels: An analysis of citi bike's incentive program. In *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies*, COMPASS '18, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450358163. doi: 10.1145/3209811.3209866. URL https://doi.org/10.1145/3209811.3209866.

[4] Claudio Contardo, Catherine Morency, and Louis-Martin Rousseau. *Balancing a dynamic public bike-sharing system*, volume 4. Cirrelt Montreal, 2012.

[5] RWR Darling, James R Norris, et al. Differential equation approximations for markov chains. *Probability surveys*, 5(1):37–79, 2008.

[6] Güneş Erdoğan, Maria Battarra, and Roberto Wolfler Calvo. An exact algorithm for the static rebalancing problem arising in bicycle sharing systems. *European Journal of Operational Research*, 245(3):667–679, 2015.

[7] Stewart N Ethier and Thomas G Kurtz. *Markov processes: characterization and convergence*, volume 282. John Wiley & Sons, 2009.

[8] Ahmadreza Faghih-Imani, Robert Cornelius Hampshire, Lavanya Marla, and Naveen Eluru. An empirical analysis of bike sharing usage and rebalancing: Evidence from barcelona and seville. 2015.

[9] Ahmadreza Faghih-Imani, Robert Hampshire, Lavanya Marla, and Naveen Eluru. An empirical analysis of bike sharing usage and rebalancing: Evidence from barcelona and seville. *Transportation Research Part A: Policy and Practice*, 97:177–191, 2017.

[10] Daniel Freund, Shane G Henderson, and David B Shmoys. Minimizing multimodular functions and allocating capacity in bike-sharing systems. In *International Conference on Integer Programming and Combinatorial Optimization*, pages 186–198. Springer, 2017.

[11] Daniel Freund, Ashkan Norouzi-Fard, Alice Paul, Carter Wang, Shane G Henderson, and David B Shmoys. Data-driven rebalancing methods for bike-share systems. In *Analytics for the Sharing Economy: Mathematics, Engineering and Business Perspectives*, pages 255–278. Springer, 2020.

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

[13] Supriyo Ghosh, Pradeep Varakantham, Yossiri Adulyasak, and Patrick Jaillet. Dynamic repositioning to reduce lost demand in bike sharing systems. *Journal of Artificial Intelligence Research*, 58:387–430, 2017.

[14] Robert C Hampshire and Lavanya Marla. An analysis of bike sharing usage: Explaining trip generation and attraction from observed demand. In *91st Annual meeting of the transportation research board, Washington, DC*, pages 12–2099, 2012.

[15] Nanjing Jian, Daniel Freund, Holly M Wiberg, and Shane G Henderson. Simulation optimization for a large-scale bike-sharing system. In *Winter Simulation Conference (WSC), 2016*, pages 602–613. IEEE, 2016.

[16] Ashish Kabra, Elena Belavina, and Karan Girotra. Bike-share systems: Accessibility and availability. 2016.

[17] Young Myoung Ko and Jamol Pender. Strong approximations for time varying infinite-server queues with non-renewal arrival and service processes. *Stochastic Models To Appear*, 2017.

[18] Rahul Nair and Elise Miller-Hooks. Equilibrium network design of shared-vehicle systems. *European Journal of Operational Research*, 235(1):47–61, 2014.

[19] Samantha Nirenberg, Andrew Daw, and Jamol Pender. The impact of queue length rounding and delayed app information on disney world queues. In *2018 Winter Simulation Conference (WSC)*, pages 3849–3860. IEEE, 2018.

[20] Sophia Novitzky, Jamol Pender, Richard Rand, and Elizabeth Wesson. Limiting the oscillations in queues with delayed information through a novel type of delay announcement. *arXiv preprint arXiv:1902.07617*, 2019.

[21] Sophia Novitzky, Jamol Pender, Richard H Rand, and Elizabeth Wesson. Nonlinear dynamics in queueing theory: Determining the size of oscillations in queues with delay. *SIAM Journal on Applied Dynamical Systems*, 18(1):279–311, 2019.

[22] Eoin O'Mahony and David B Shmoys. Data analysis and optimization for (citi) bike sharing. In *AAAI*, pages 687–694, 2015.

[23] Eoin Daniel O'Mahony. *Smarter tools for (Citi) bike sharing*. PhD thesis, Cornell University, 2015.

[24] Jamol Pender, Richard H Rand, and Elizabeth Wesson. Queues with choice via delay differential equations. *International Journal of Bifurcation and Chaos*, 27(04):1730016, 2017.

[25] Jamol Pender, Richard H Rand, and Elizabeth Wesson. An analysis of queues with delayed information and time-varying arrival rates. *Nonlinear Dynamics*, 91(4):2411–2427, 2018.

[26] Julius Pfrommer, Joseph Warrington, Georg Schildbach, and Manfred Morari. Dynamic vehicle redistribution and online price incentives in shared mobility systems. *IEEE Transactions on Intelligent Transportation Systems*, 15(4):1567–1578, 2014.

[27] Tal Raviv, Michal Tzur, and Iris A Forma. Static repositioning in a bike-sharing system: models and solution approaches. *EURO Journal on Transportation and Logistics*, 2(3): 187–229, 2013.

[28] Jasper Schuijbroek, Robert C Hampshire, and W-J Van Hoeve. Inventory rebalancing and vehicle routing in bike sharing systems. *European Journal of Operational Research*, 257(3):992–1004, 2017.

[29] Adish Singla, Marco Santoni, Gábor Bartók, Pratik Mukerji, Moritz Meenen, and Andreas Krause. Incentivizing users for balancing bike sharing systems. In *AAAI*, pages 723–729, 2015.

[30] Shuang Tao and Jamol Pender. A stochastic analysis of bike sharing systems. *arXiv preprint arXiv:1708.08052*, 2017.

[31] Ariel Waserhole and Vincent Jost. Pricing in vehicle sharing systems: Optimization in queuing networks with product forms. *EURO Journal on Transportation and Logistics*, 5(3):293–320, 2016.

[32] Ariel Waserhole, Vincent Jost, and Nadia Brauner. Pricing techniques for self regulation in vehicle sharing systems. *Electronic Notes in Discrete Mathematics*, 41:149–156, 2013.