A STOCHASTIC ANALYSIS OF BIKE-SHARING SYSTEMS

SHUANG TAO AND JAMOL PENDER ⓘD

School of Operations Research and Information Engineering, Cornell University, 293 Rhodes Hall, Ithaca, NY 14853, USA

E-mail: st754@cornell.edu; jjp274@cornell.edu

As more people move back into densely populated cities, bike sharing is emerging as an important mode of urban mobility. In a typical bike-sharing system (BSS), riders arrive at a station and take a bike if it is available. After retrieving a bike, they ride it for a while, then return it to a station near their final destinations. Since space is limited in cities, each station has a finite capacity of docks, which cannot hold more bikes than its capacity. In this paper, we study BSSs with stations having a finite capacity. By an appropriate scaling of our stochastic model, we prove a mean-field limit and a central limit theorem for an empirical process of the number of stations with k bikes. The mean-field limit and the central limit theorem provide insight on the mean, variance, and sample path dynamics of large-scale BSSs. We also leverage our results to estimate confidence intervals for various performance measures such as the proportion of empty stations, the proportion of full stations, and the number of bikes in circulation. These performance measures have the potential to inform the operations and design of future BSSs.

Keywords: applied probability, operations research, probabilistic networks, queueing theory, stochastic modeling

1. INTRODUCTION

Bike sharing is an emerging mode of eco-friendly transportation that has launched in over 400 cities around the world [19]. In the United States, we are witnessing a transition where more people are deciding to live in large and densely populated cities. As more people transition from the sprawling suburbs into densely populated cities, bike-sharing programs will continue to grow in popularity since they provide easy transportation for citizens of these large cities. As more people use these bike-sharing systems (BSSs), less people will drive motor vehicles on the road. This reduction of vehicles on the road due to BSS has the potential to also reduce the growing traffic congestion in these growing cities. BSS also promote healthy living as biking is a great form of exercise. They are environmentally friendly and they have the potential to reduce carbon emissions if operated correctly and efficiently; see, for example, Hampshire and Marla [13], Nair and Miller-Hooks [19], Nair *et al.* [21], Nair and Miller-Hooks [20], O'Mahony and Shmoys [25], O'Mahony [24], Jian *et al.* [14], Freund *et al.* [9], and their references within for more information on BSS.
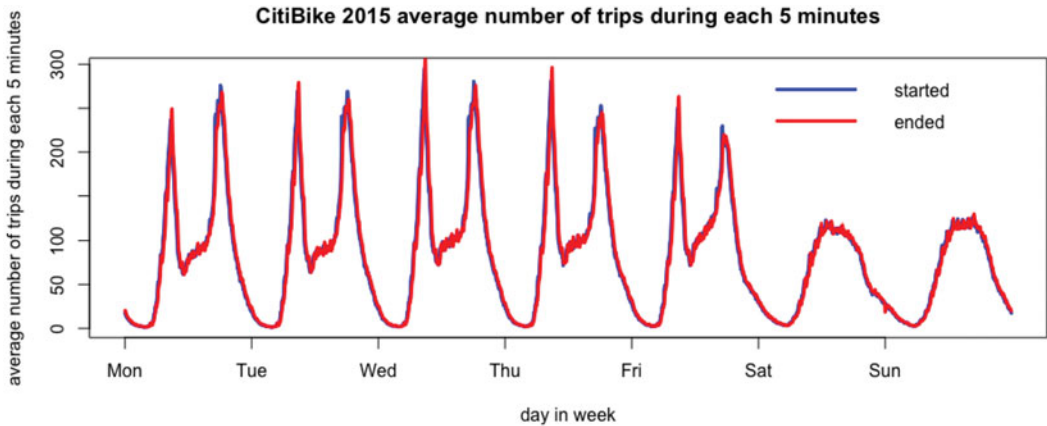
FIGURE **1.** Citi Bike average number of trips during each 5 min (January 1–December 31, 2015). The red line represents the average number of trips that started during each 5 min. The blue line represents the average number of trips that ended during each 5 min.

For a typical system, riders simply arrive at a station and select a bike if there is one available for them to take. If there is no bike available for the rider, the rider will leave the system. Otherwise, the rider will take a bike and ride it for a while before returning it to another station near their final destination, if there is available space. If no space is available, the rider must find a nearby station to return the bike. If there were an infinite supply of docks to store the bikes, then our model would be reduced to a network of infinite server queues, which is more tractable to analyze. However, since the number of bike docks have finite capacity, the model become less tractable especially for systems with a large number of stations.

Much of the complexity inherent in BSS lies in the scarcity of resources to move all riders around each city at all times of the day. Riders can encounter the scarcity of resources in two fundamental ways. First, a rider can encounter insufficient resources by finding an empty station with no bikes when a rider needs one. Secondly, a rider can find a station full with bikes when attempting to return a bike. Thus, from a managerial point of view, having stations with no bikes or too many bikes are both problematic for riders. There are several reasons why bike stations either have no or too many bikes. One main reason why stations might have too many bikes or too few bikes is that the system is highly inhomogeneous. Not only is the arrival rate a non-constant function of the time of day (see Figure 1) but also riders do not evenly distribute themselves amongst the available stations. For example, many stations that are located in residential areas have fewer bikes available for riding, as many riders take bikes to more commercial areas. Another example that illustrates the inhomogeneous dynamics is that riders tend to take bikes from up-hill stations to go to down-hill stations; however, very few riders take bikes from down-hill stations to go up-hill. Thus, as more bikes flow from residential to commercial areas during rush hours, or up-hill to down-hill stations, this causes the system to be more imbalanced over time.

Since BSS are quite complex, researchers have been inspired to study these systems in great depth. The subsequent analysis of BSS has generated many insights on these systems, especially for rebalancing the fleet of bikes. Although there is a large community that studies these systems, few analytical models have been proposed, especially stochastic analytical models. This is primarily because the stochastic models for BSS are often very complex

and are rather intractable to analyze without making strong assumptions. Nevertheless, the analysis of such stochastic and mathematical models could provide insights on the behavior of these BSS and how to manage them effectively. In fact, a deeper analysis of stochastic models for such systems could help researchers understand the impact of different incentive algorithms for taking or returning bikes, which can lead to significant improvements in the overall system performance.

Most BSS can be viewed as closed queueing networks. One of the first papers to model the BSS as a closed queueing network is by George and Xia [11]. However, in the bike-sharing context, the customers are replaced by bikes. The number of bikes, also known as the fleet size, is fixed and remains constant. In this model, the bikes can go to two types of stations. The first type of station is a single-server queue where the service times are the user inter-arrival times to this station. The second type of station is an infinite-server queue where the service times are the trip times on a route from station i to station j. The main drawback of the model by George and Xia [11] is that it is based on infinite capacity queues. This means that the model does not take into account the finite capacity of the stations and the related strategies of the users to return their bikes. To overcome this major drawback, the model proposed in this paper allow the finite capacity at stations. Thus, we are able to model the real system where customers are blocked from returning bikes to the stations that is nearest to their destination.

In our model, we model the BSS as state-dependent $M/M/1/K_i$ queueing networks. When joining a saturated single-server queue, the user reattempts in another queue, after a time with the same distribution as the trip time, until he returns his bike. Although this model seems to model the behavior of the network, it is not practical since it scales with the number of stations. Thus, we follow an approach developed by Fricker *et al.* [10] to study the bike-sharing network's empirical process instead. The empirical process still allows us to derive important performance measures of the original system; however, it scales with the maximum station capacity and not the number of stations, which is more practical for large networks like Citi Bike.

1.1. Main Contributions of Paper

The contributions of this work can be summarized as follows:

- We construct a stochastic bike-sharing queueing model that incorporates the finite capacity of stations. Since our model is difficult to analyze for a large number of stations, we propose to analyze an empirical process that describes the proportion of stations that have a certain number of bikes.
- We prove a mean-field limit and a central limit theorem for our stochastic bike-sharing empirical process, showing that the mean-field limit and the variance of the empirical process can be described by a system of $\frac{1}{2}(K + 4)(K + 1)$ differential equations, where K is the maximum station capacity.
- Using the mean-field limit of the empirical process, we are able to approximate the mean proportion of empty and full stations. Furthermore, with the central limit theorem of the empirical process, we are able to construct confidence intervals around the mean-field limit for the same performance measures.
- We compare the mean-field limit and the central limit theorem to a simulation and show that the differential equations approximate the mean and variance of the empirical process extremely well.

1.2. Organization of Paper

The remainder of this paper is organized as follows. Section 2 provides a brief history of bike-sharing programs and a literature review on the research streams concerning BSS. Section 3 introduces our bike-sharing model and notation of the paper. In Section 4, we derive and prove the mean-field limit of the empirical measure process of the distribution of stations with different bikes and utilization rate. In Section 5, we derive the diffusion limit and prove a functional CLT for our model. We show that the diffusion process is a centered Gaussian OU process and we also obtain a closed-form expression of the diffusion limit process. In Section 6, we extend our analysis in Section 3 to a broader case with non-uniform routing probabilities and capacities and derive the mean-field limit and diffusion limit in this extended case. In Section 7, we discuss the simulation results of our model, with both stationary and non-stationary arrival processes. We also give a comparison between real Citi Bike empirical measure and simulated ones using our model to show how well our model is in capturing reality. Finally, in Section 9, we give concluding remarks and provide some future directions of research that we intend to pursue later.

2. HISTORY AND LITERATURE REVIEW

The literature that focuses on the analysis and operations of BSS is increasing rapidly. Early research that studied the history of BSS includes Shaheen *et al.* [35], Hampshire and Marla [13], Nair *et al.* [21], Schuijbroek *et al.* [34], and DeMaio [6]. These papers provide a history of bike sharing and how it has evolved over time. The beginning of bike sharing can be traced back to the first generation of *white bikes* (or free bikes) in Amsterdam, The Netherlands, as early in 1965. However, this first generation of BSS failed due to a large amount of bike theft. The launch of Bycyklen in Copenhagen in 1995 marked the second generation of BSS. Bycyklen was the first BSS to implement docking stations and coin-deposit systems to unlock bikes. These coin-deposit systems helped with the bike theft problem and, thus, made the BSS more reliable. However, even with these improvements, the Bycyklen could not eliminate bike thefts mainly due to the fact that customer still remained anonymous and there were no time limits on how long a customer could use a bike.

The failures of the second generation of BSS inspired the present-day or the third generation of BSS by combining docks with information technology. These new systems incorporate information technology for bicycle reservations, pickups, and drop-offs. This new technology has enabled many BSS to keep better track of bicycles and the users that use them, thus eliminating virtually all bike theft. One example of a third-generation BSS is the Paris' bike-sharing program Velib. Velib was launched in Paris in July 2007 and has emerged as the most prominent example of a successful bike-sharing program in the modern world. As a result of the success of Velib, many cities like New York City (Citi Bike), Chicago (Divvy), and even Ithaca (Big Red Bikes) have implemented large-scale BSS and bike sharing has become a widely used form of transportation in these cities. For the interested reader, Laporte *et al.* [17] provide a comprehensive survey of the vehicle/bike-sharing literature.

Rebalancing is currently the biggest stream of research concerning BSS. In rebalancing operations, there are two methods of rebalancing: (1) deploying a truck fleet or (2) providing user incentives, where deploying a truck fleet is often referred to as bike repositioning. Both methods involve static and dynamic cases. Static repositioning usually is moving bikes during the night when traffic flow is low, while dynamic repositioning is moving bikes during the day based on the current state of the system. Most research on this area focuses on the static case, partly because it is easier to model and also because the impact of repositioning

is more important during the night [14]. Raviv *et al.* [33] is one of the first papers to study static repositioning of BSS, using mixed integer linear programming by maximizing customer demand satisfaction. Benchimol *et al.* [2] consider a similar problem, where a single truck repositions bikes to bring the inventory of each station to a predetermined value. However, their objective is to minimize the routing cost as opposed to maximizing customer satisfaction. In the case of dynamic repositioning, Chemla *et al.* [3] and Pfrommer *et al.* [32] consider the case when the trucks respond in real time to the current state of the system. However, Contardo *et al.* [4] and Ghosh *et al.* [12] consider the situation where the time-dependent demand is known *a priori* and the rebalancing operations are computed in an off-line fashion. Yet, none of these papers really explore stochastic dynamics and they for the most part mainly exploit optimization techniques to tackle the problem.

Unlike the rebalancing literature, our paper falls into the performance analysis literature with an emphasis in supply analysis. We focus on analyzing the most salient performance measures such as the mean, variance, covariance, and sample path dynamics of the bike distributions in a large-scale BSS. There is not much literature that explores the fluctuations of BSS around the mean-field limit. In this paper, we prove a mean-field limit and a functional central limit theorem under some smoothness conditions and show that the diffusion limit is characterized by a multi-dimensional Ornstein–Uhlenbeck (OU) process. The functional central limit theorem not only gives us information about the sample path fluctuations of the queue length process, but it also allows us to construct approximate confidence intervals for various performance measures such as the proportion of empty stations, the proportion of full stations, and the mean number of bikes in circulation. Unlike the previous literature, our paper also considers non-stationary arrivals to stations, which is much more realistic given the user patterns we observe in the historical data from CitiBike. In Figure 1, we plot the empirical mean of the number of trips (5 min intervals) during the week for the time period January 1–December 31, 2015. We observe from Figure 1 that the arrival rate is non-stationary and clearly reflects the morning rush and evening rush during the peak times. Thus, analyzing the non-stationary dynamics is crucial for understanding the impact of system inhomogeneity since it can provide useful guidelines for rebalancing operations.

3. BIKE-SHARING QUEUEING MODEL

In this section, we construct a Markovian bike-sharing queueing model where customers can pickup and drop-off bikes at each station if there is available capacity. Figure 2 provides an illustration of a typical Citi Bike station in New York City (NYC). As one can see in Figure 2, the bikes are attached to docks and the number of docks is finite with roughly 40 docks. Figure 3 shows a map of Citi Bike stations, the nation's largest bike-sharing program, with over 10,000 bikes and 600 stations across Manhattan, Brooklyn, Queens, and Jersey City. Citi Bike was designed for quick, affordable, and convenient trips and has become an essential part of the transportation infrastructure in NYC.

Motivated by the Citi Bike BSS, we consider a BSS 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 station i. When a customer arrives at a station, if there is no bikes available, they will then leave the system and are immediately blocked and lost. Otherwise, the customer will take a bike and ride 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 station i 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

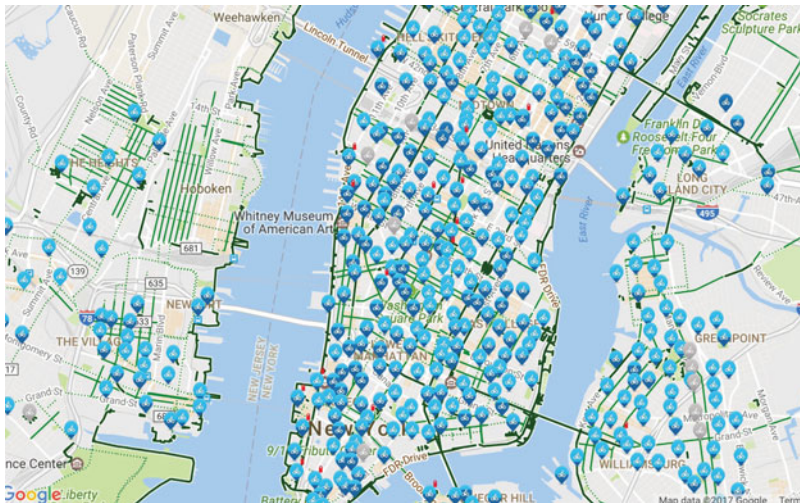FIGURE **2.** Illustration of a typical Citi Bike station.

FIGURE **3.** Citi Bike stations map.

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 so that they understand the notation that we will use throughout the paper.

To avoid cumbersome notation, throughout Sections 3–5, we *assume without loss of generality* that the service rate μ is equal to 1 and that the routing probability from each station is uniform, that is, $P_i = 1/N$. We also assume that capacities across all stations are equal, that is, $K_i = K$, for $i = 1, \ldots, N$. With our notation in hand, we are ready to

TABLE **1.** Notation

N	Number of stations
M	Total 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
$R_i = \mu P_i / \lambda_i$	Utilization at station i
$r_i = R_i / \max_i R_i$	Relative utilization at station i
γ	Average number of bikes at each station

develop our stochastic model for our bike-sharing network. At first glance, these assumptions seem restrictive; however, we explain in Section 6.1 how our model can be extended seamlessly to more complex settings with non-uniform routing probabilities and capacities. We should mention, however, the extension to inter-station transition probabilities, that is, probabilities that depends on the departing station as well as the returning station are non-Markovian models. In Section 6.2, we discuss more in detail how one can extend the state space, by tracking the number of in-transit bikes coming from each station, to make the queueing model Markovian; however, this extended model makes the problem even more high-dimensional and adds difficulty to the analysis.

We define $\mathbf{X}(t) = (X_1(t), \ldots, X_N(t))$, 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 $\mu P_i (M - \sum_{k=1}^{N} X_k(t)) \mathbf{1}\{X_i(t) > K_i\}$, and the rate of retrieving bikes at station i is equal to $\lambda_i \mathbf{1}\{X_i(t) < 0\}$. Using these rates, we can construct the functional forward equations for our stochastic bike-sharing model. This construction is given below in the following proposition:

PROPOSITION 3.1: *For any integrable function* $f : \mathbb{Z}_+^N \to \mathbb{R}$, *the CTMC* $\mathbf{X}(t)$ *satisfy the following functional forward equation:*

$$\dot{\mathbb{E}}[f(\mathbf{X}(t))|\mathbf{X}(0) = \mathbf{x}] \equiv \frac{d}{dt}\mathbb{E}[f(\mathbf{X}(t))|\mathbf{X}(0) = \mathbf{x}]$$

$$= \sum_{i=1}^{N} \mathbb{E}[(f(\mathbf{X}(t) - \mathbf{1}_i) - f(\mathbf{X}(t))\lambda_i \mathbf{1}_{\{X_i(t)<0\}}]$$

$$+ \sum_{i=1}^{N} \mathbb{E}\left[(f(\mathbf{X}(t) + \mathbf{1}_i) - f(\mathbf{X}(t))\,\mu P_i \left(M - \sum_{k=1}^{N} X_k(t)\right) \mathbf{1}_{\{X_i(t)>K_i\}}\right]. \qquad \textbf{(3.1)}$$

PROOF: The proof is found in the Appendix. ∎

COROLLARY 3.2: *The time derivatives of the mean, variance, and covariance of* $\mathbf{X}(t)$ *are given by*

$$\dot{\mathbb{E}}[X_i(t)] = \mathbb{E}\left[\mu P_i \left(M - \sum_{k=1}^{N} X_k(t)\right) \mathbf{1}\{X_i(t) > K_i\}\right] - \lambda_i \mathbb{P}[X_i(t) < 0], \qquad \textbf{(3.2)}$$

$$\dot{\mathrm{Var}}[X_i(t)] = \dot{\mathbb{E}}[X_i^2(t)] - 2\dot{\mathbb{E}}[X_i(t)]\mathbb{E}[X_i(t)]$$

$$= 2\mathrm{Cov}\left[X_i(t), \mu P_i\left(M - \sum_{k=1}^{N} X_k(t)\right)\mathbf{1}\{X_i(t) > K_i\}\right]$$

$$- 2\mathrm{Cov}[X_i(t), \lambda_i \mathbf{1}\{X_i(t) < 0\}]$$

$$+ \mathbb{E}\left[\mu P_i\left(M - \sum_{k=1}^{N} X_k(t)\right)\mathbf{1}\{X_i(t) > K_i\}\right] + \lambda_i \mathbb{P}[X_i(t) < 0], \qquad \textbf{(3.3)}$$

$$\dot{\mathrm{Cov}}[X_i(t), X_j(t)] = \mathrm{Cov}\left[\mu P_j\left(M - \sum_{k=1}^{N} X_k(t)\right)\mathbf{1}_{\{X_j(t) > K_j\}} - \lambda_j \mathbf{1}_{\{X_j(t) < 0\}}, X_i(t)\right]$$

$$+ \mathrm{Cov}\left[\mu P_i\left(M - \sum_{k=1}^{N} X_k(t)\right)\mathbf{1}_{\{X_i(t) > K_i\}} - \lambda_i \mathbf{1}_{\{X_i(t) < 0\}}, X_j(t)\right],$$

$$\textbf{(3.4)}$$

for $i, j = 1, \ldots, N$ and $i \neq j$.

3.1. Intractability of Individual Stations Model

Although the functional forward equations given in Corollary 3.2 describe the exact dynamics of the mean, variance, and covariance of the BSS, 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 BSS 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. Work by Massey and Pender [18], Pender [26], Pender [27], Engblom and Pender [7], and Pender [28] could yield useful and accurate closure approximations for making the system closed. Moreover, with the exception of Engblom and Pender [7], there are no error bounds on the accuracy of various closure approximations. Thus, it is not clear how well the closure approximations would perform over a variety of parameter settings. 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 (around 180,900 equations in the Citi Bike case) that would need to be numerically integrated. This is very computationally expensive, and thus, we must take a different approach to analyze our BSS.

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 BSS to our advantage and have a finite dimensional CTMC for the empirical measure process.

3.2. An Empirical Measure Model

Following the model of Fricker *et al.* [10], we construct an empirical measure process that counts the proportions of stations with n bikes and utilization r. This empirical measure process is given by the following equation:

$$Y_t^N(r, n) = \frac{1}{N}\sum_{i=1}^{N}\mathbf{1}\{r_i^N = r, X_i^N(t) = n\}. \qquad \textbf{(3.5)}$$

We further define that $Y_t^N(n) = \sum_r Y_t^N(r,n)$. By observing the empirical measure process, we notice that $Y_t^N = (Y_t^N(0), \ldots, Y_t^N(K)) \in [0,1]^{K+1}$. Thus, for our empirical measure process, we only need to solve $O(K^2)$ differential equations for understanding the mean and variance dynamics of the BSS, where $K \ll N$. More importantly, the empirical measure will also allow us to obtain salient performance measures such as $Y_t^N(0)$ (the proportion of stations with no bikes), $Y_t^N(K)$ (the proportion of stations that are full of bikes), $M - \sum_{j=0}^{K} j \cdot Y_t^N(j)N$ (the number of bikes in circulation), and among others.

Conditioning on $Y_t^N(r,n) = y(r,n)$ and given our assumptions that $\mu = 1$ and $P_i = 1/N$, the transition rates of y are specified as follows:

When a customer arrives 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

$$
\begin{aligned}
Q^N\left(y, y + \frac{1}{N}(\mathbf{1}_{(r,n-1)} - \mathbf{1}_{(r,n)})\right) &= y(r,n)\lambda_r N \mathbf{1}_{n<0} \\
&= y(r,n)\frac{\mu P_i}{R}N\mathbf{1}_{n<0} \\
&= y(r,n)\frac{1}{NR}N\mathbf{1}_{n<0} \\
&= \frac{y(r,n)}{rR_{\max}^N}\mathbf{1}_{n<0}.
\end{aligned}
\tag{3.6}
$$

When a customer returns 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

$$
\begin{aligned}
Q^N\left(y, y + \frac{1}{N}(\mathbf{1}_{(r,n+1)} - \mathbf{1}_{(r,n)})\right) &= y(r,n) \cdot \mu \cdot \left(M - \sum_{n'}\sum_{r'} n'y(r',n')N\right)\mathbf{1}_{n>K} \\
&= y(r,n)N\left(\frac{M}{N} - \sum_{n'}\sum_{r'} n'y(r',n')\right)\mathbf{1}_{n>K}.
\end{aligned}
\tag{3.7}
$$

Similarly, we have the functional forward equations for $Y_t^N(r,n)$.

PROPOSITION 3.3: *For any integrable function $f : [0,1]^{K+1} \to \mathbb{R}$, $Y_t^N(r) = (Y_t^N(r,0), \ldots, Y_t^N(r,K))$ satisfies the following functional forward equation:*

$$
\begin{aligned}
\dot{\mathbb{E}}&(f(Y_t^N(r))|Y_0^N(r) = y_0(r)] \\
&= \sum_{n=0}^{K} \mathbb{E}\left[\left(f\left(Y_t^N(r) + \frac{1}{N}(\mathbf{1}_{r,n-1} - \mathbf{1}_{r,n})\right) - f(Y_t^N(r))\right)\frac{Y_t^N(r,n)}{rR_{\max}}\mathbf{1}_{n<0}\right] \\
&\quad + \sum_{n=0}^{K} \mathbb{E}\left[\left(f\left(Y_t^N(r) + \frac{1}{N}(\mathbf{1}_{r,n+1} - \mathbf{1}_{r,n})\right) - f(Y_t^N(r))\right)\right. \\
&\quad \left.\times Y_t^N(r,n)N\left(\frac{M}{N} - \sum_{n'}\sum_{r'} n'Y_t^N(r',n')\right)\mathbf{1}_{n>K}\right].
\end{aligned}
\tag{3.8}
$$

PROOF: The proof is similar to the proof of Proposition 3.1. ∎

COROLLARY 3.4: *The time derivative of the mean of* $Y_t^N(r, n)$ *is given by*

$$\dot{\mathbb{E}}[Y_t^N(r,n)] = \mathbb{E}\left[\frac{1}{rNR_{\max}^N}\left(Y_t^N(r,n+1)\mathbf{1}_{n>K} - Y_t^N(r,n)\mathbf{1}_{n<0}\right)\right]$$

$$+ \mathbb{E}\left[\left(\frac{M}{N} - \sum_{n'}\sum_{r'}n'Y_t^N(r',n')\right)\left(Y_t^N(r,n-1)\mathbf{1}_{n<0} - Y_t^N(r,n)\mathbf{1}_{n>K}\right)\right]. \quad \textbf{(3.9)}$$

for $n = 0, \ldots, K$. *Denote* $\Sigma_{i,j}^N(r,t) = \mathrm{Cov}[Y_t^N(r,i), Y_t^N(r,j)]$. *When* $i = j$, *the time deriva-tive of the variance term* $\mathrm{Var}[Y_t^N(r,i)]$ *is given by*

$$\dot{\mathrm{Var}}[Y_t^N(r,i)] = \dot{\mathbb{E}}[Y_t^N(r,i)^2] - 2\dot{\mathbb{E}}[Y_t^N(r,i)]\mathbb{E}[Y_t^N(r,i)]$$

$$= \frac{2}{rNR_{\max}^N}(\Sigma_{i,i+1}^N(r,t)\mathbf{1}_{i>K} - \Sigma_{i,i}^N(r,t)\mathbf{1}_{i<0})$$

$$+ \frac{2M}{N}\left(\Sigma_{i,i-1}^N(r,t)\mathbf{1}_{i<0} - \Sigma_{i,i}^N(r,t)\mathbf{1}_{i>K}\right)$$

$$- 2\sum_{n'}\sum_{r'}n'(\mathrm{Cov}[Y_t^N(r,i), Y_t^N(r',n')Y_t^N(r,i-1)]\mathbf{1}_{i<0}$$

$$- \mathrm{Cov}[Y_t^N(r,i), Y_t^N(r',n')Y_t^N(r,i)]\mathbf{1}_{i>K})$$

$$+ \frac{1}{rN^2R_{\max}^N}(\mathbb{E}[Y_t^N(r,i+1)]\mathbf{1}_{i>K} + \mathbb{E}[Y_t^N(r,i)]\mathbf{1}_{i<0})$$

$$+ \frac{M}{N^2}(\mathbb{E}[Y_t^N(r,i-1)]\mathbf{1}_{i<0} + \mathbb{E}[Y_t^N(r,i)]\mathbf{1}_{i>K})$$

$$- \frac{1}{N}\sum_{n'}\sum_{r'}n'(\mathbb{E}[Y_t^N(r',n')Y_t^N(r,i-1)]\mathbf{1}_{i<0}$$

$$+ \mathbb{E}[Y_t^N(r',n')Y_t^N(r,i)]\mathbf{1}_{i>K}). \quad \textbf{(3.10)}$$

When $|i - j| < 1$, *the time derivative of the covariance term* $\mathrm{Cov}[Y_t^N(r,i), Y_t^N(r,j)]$ *is given by*

$$\dot{\mathrm{Cov}}[Y_t^N(r,i), Y_t^N(r,j)]$$

$$= \dot{\mathbb{E}}[Y_t^N(r,i)Y_t^N(r,j)] - \dot{\mathbb{E}}[Y_t^N(r,i)]\mathbb{E}[Y_t^N(r,j)] - \dot{\mathbb{E}}[Y_t^N(r,j)]\mathbb{E}[Y_t^N(r,i)]$$

$$= \frac{1}{rNR_{\max}^N}[\Sigma_{i+1,j}^N(r,t)\mathbf{1}_{i>K} + \Sigma_{i,j+1}^N(r,t)\mathbf{1}_{j>K} - \Sigma_{i,j}^N(r,t)(\mathbf{1}_{j<0} + \mathbf{1}_{i<0})]$$

$$+ \frac{M}{N}[\Sigma_{i-1,j}^N(r,t)\mathbf{1}_{i<0} + \Sigma_{i,j-1}^N(r,t)\mathbf{1}_{j<0} - \Sigma_{i,j}^N(r,t)(\mathbf{1}_{i>K} + \mathbf{1}_{j>K})]$$

$$- \sum_{n'}\sum_{r'}n'(\mathrm{Cov}[Y_t^N(r,j), Y_t^N(r',n')Y_t^N(r,i-1)]\mathbf{1}_{i<0}$$

$$- \mathrm{Cov}[Y_t^N(r,j), Y_t^N(r',n')Y_t^N(r,i)]\mathbf{1}_{i>K})$$

$$- \sum_{n'}\sum_{r'}n'(\mathrm{Cov}[Y_t^N(r,i), Y_t^N(r',n')Y_t^N(r,j-1)]\mathbf{1}_{j<0}$$

$$- \mathrm{Cov}[Y_t^N(r,i), Y_t^N(r',n')Y_t^N(r,j)]\mathbf{1}_{j>K}) \quad \textbf{(3.11)}$$

and when $j = i + 1$, the time derivative of the covariance term $\mathrm{Cov}[Y_t^N(r, i), Y_t^N(r, i + 1)]$ is given by

$$
\begin{aligned}
\dot{\mathrm{Cov}}&[Y_t^N(r, i), Y_t^N(r, i + 1)] \\
&= \dot{\mathbb{E}}[Y_t^N(r, i) Y_t^N(r, i + 1)] - \dot{\mathbb{E}}[Y_t^N(r, i)]\mathbb{E}[Y_t^N(r, i + 1)] - \dot{\mathbb{E}}[Y_t^N(r, i + 1)]\mathbb{E}[Y_t^N(r, i)] \\
&= \frac{1}{rNR_{\max}^N}[\Sigma_{i+1,i+1}^N(r, t) + \Sigma_{i,i+2}^N(r, t)\mathbf{1}_{i>K-1} - \Sigma_{i,i+1}^N(r, t)(1 + \mathbf{1}_{i<0})] \\
&\quad + \frac{M}{N}[\Sigma_{i-1,i+1}^N(r, t)\mathbf{1}_{i<0} + \Sigma_{i,i}^N(r, t) - \Sigma_{i,i+1}^N(r, t)(1 + \mathbf{1}_{i>K-1})] \\
&\quad - \sum_{n'}\sum_{r'} n'(\mathrm{Cov}[Y_t^N(r, i + 1), Y_t^N(r', n')Y_t^N(r, i - 1)]\mathbf{1}_{i<0} \\
&\quad - \mathrm{Cov}[Y_t^N(r, i + 1), Y_t^N(r', n')Y_t^N(r, i)]) \\
&\quad - \sum_{n'}\sum_{r'} n'(\mathrm{Cov}[Y_t^N(r, i), Y_t^N(r', n')Y_t^N(r, i)] \\
&\quad - \mathrm{Cov}[Y_t^N(r, i), Y_t^N(r', n')Y_t^N(r, i + 1)]\mathbf{1}_{i>K-1} \\
&\quad - \frac{\mathbb{E}[Y_t^N(r, i + 1)]}{rN^2 R_{\max}} - \frac{M}{N^2}\mathbb{E}[Y_t^N(r, i)] + \frac{1}{N}\sum_{n'}\sum_{r'} n'\mathbb{E}[Y_t^N(r', n')Y_t^N(r, i)]. \qquad \textbf{(3.12)}
\end{aligned}
$$

PROOF: The time derivatives of $\mathbb{E}[Y_t^N(r, i)]$ and $\mathrm{Var}[Y_t^N(r, i)]$ come directly from applying Proposition 3.3 with $f(Y_t(r)) = Y_t(r, i), Y_t^2(r, i)$, respectively. For the covariance term, we let $f(Y_t^N(r)) = f(Y_t^N(r, i), Y_t^N(r, j))$ to prove the result. ■

Although we have equations for the moments of the empirical process and individual stations, it is still difficult to analyze them and gain insights from them directly. One reason is that even though we have reduced the dimensionality of the analysis significantly, we have not removed the non-closure property of the differential equations. Thus, we need to develop a new approach that will allow us to get around this complication. The method that we choose to use is asymptotic analysis and will be described in more details in the sequel. However, before we get to the asymptotic analysis, we state some technicalities about weak convergence.

3.3. Preliminaries of Weak Convergence

Following Ko and Pender [15], we assume that all random variables in this paper are defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Moreover, for all positive integers k, we let $\mathcal{D}([0, \infty), \mathbb{R}^k)$ be the space of right-continuous functions with left limits (RCLL) in $\mathbb{R}^k$ that have a time domain in $[0, \infty)$. As is usual, we endow the space $\mathcal{D}([0, \infty), \mathbb{R}^k)$ with the usual Skorokhod J_1 topology, and let M^k be defined as the Borel σ-algebra associated with the J_1 topology. We also assume that all stochastic processes are measurable functions from our common probability space $(\Omega, \mathcal{F}, \mathbb{P})$ into $(\mathcal{D}([0, \infty), \mathbb{R}^k), M^k)$. Thus, if $\{\zeta\}_{n=1}^\infty$ is a sequence of stochastic processes, then the notation $\zeta^n \to \zeta$ implies that the probability measures that are induced by the ζ^n's on the space $(\mathcal{D}([0, \infty), \mathbb{R}^k), M^k)$ converge weakly to the probability measure on the space $(\mathcal{D}([0, \infty), \mathbb{R}^k), M^k)$ induced by ζ. For any $x \in (\mathcal{D}([0, \infty), \mathbb{R}^k), M^k)$ and any $T < 0$, we define

$$
|x|_T \equiv \sup_{0 \le t \le T} \max_{i=1,2,\ldots,k} |x_i(t)| \qquad \textbf{(3.13)}
$$

and note that ζ^n converges almost surely to a continuous limit process ζ in the J_1 topology if and only if

$$|\zeta^n - \zeta|_T \to 0 \quad a.s. \tag{3.14}$$

for every $T < 0$.

4. MEAN-FIELD LIMIT

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 obtain new insights on average system dynamics, when 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.

First, we state the important assumptions that will be used throughout the paper, to ensure the existence of a mean-field limit of our model.

ASSUMPTION 4.1: *There exists a probability measure $I(r, k)$ on $]0, 1] \times \mathbb{N}$ with finite support and $\Lambda < 0$ such that, as N tends to infinity, we have*

(i) $(1/N) \sum_{i=1}^{N} \mathbf{1}_{(r_i^N, K_i^N)} \Rightarrow I(r, k)$,
(ii) $NR_{\max}^N \to \Lambda^{-1}$,
(iii) $M/N \to \gamma$.

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

THEOREM 4.2 : *(Functional Law of Large Numbers): Let $|\cdot|$ denotes the Euclidean norm in $\mathbb{R}^{K+1}$. Under Assumption 4.1, suppose that $Y_0^N \xrightarrow{p} y_0$, then we have for any $\epsilon < 0$ and $t_0 < 0$,*

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

Here, $y_t = (y_t(0), \ldots, y_t(K))$, where $y_t(k) = \int_0^1 dy_t(r, k)$ for $k = 0, \ldots, K$. And y_t is the unique solution to the following differential equation starting at y_0 :

$$\dot{y}_t = b(y_t), \tag{4.1}$$

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

$$b(y_t) = \iint_{]0,1] \times [0,...,K]} \left[\frac{\Lambda}{r} (\mathbf{1}_{(r,n-1)} - \mathbf{1}_{(r,n)}) \mathbf{1}_{n<0} \right.$$
$$\left. + \left(\gamma - \sum_n \int_0^1 n \, dy_t(r, n) \right) (\mathbf{1}_{(r,n+1)} - \mathbf{1}_{(r,n)}) \mathbf{1}_{n>K} \right] dy_t(r, n), \tag{4.2}$$

or componentwise

$$b(y_t)(0) = \underbrace{- \int_0^1 \left(\gamma - \sum_n \int_0^1 n \, dy_t(r,n) \right) dy_t(r,0)}_{\text{return a bike to a no}-\text{bike station}} + \underbrace{\int_0^1 \frac{\Lambda}{r} dy_t(r,1)}_{\text{retrieve a bike from a 1}-\text{bike station}} ,$$

$$b(y_t)(k) = \underbrace{\int_0^1 \frac{\Lambda}{r} dy_t(r,k+1)}_{\text{retrieve a bike from a } k+1-\text{bike station}} - \underbrace{\int_0^1 \left(\frac{\Lambda}{r} + \gamma - \sum_n \int_0^1 n \, dy_t(r,n) \right) dy_t(r,k)}_{\text{retrieve and return a bike to a } k-\text{bike station}}$$

$$+ \underbrace{\int_0^1 \left(\gamma - \sum_n \int_0^1 n \, dy_t(r,n) \right) dy_t(r,k-1)}_{\text{return a bike to a } k-1-\text{bike station}},$$

for $k = 1, \ldots, K-1$, *and*

$$b(y_t)(K) = \underbrace{- \int_0^1 \frac{\Lambda}{r} dy_t(r,K)}_{\text{retrieve a bike from a } K-\text{bike station}} + \underbrace{\int_0^1 \left(\gamma - \sum_n \int_0^1 n \, dy_t(r,n) \right) dy_t(r,K-1)}_{\text{return a bike to a } K-1-\text{bike station}}.$$

PROOF: A similar theorem is given in the paper of Fricker *et al.* [10]; however, a proof is not given in their work. Thus, to make our paper self-contained, we provide a full proof of the mean-field limit for the convenience of the reader as it is essential for our future results. Our proof exploits Doob's inequality for martingales and Gronwall's lemma. Moreover, we use Proposition 4.2, Proposition 4.3, and Proposition 4.4 in the proof, and they are stated after the proof of Theorem 6.11.

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

$$Y_t^N = \underbrace{Y_0^N}_{\text{initial condition}} + \underbrace{M_t^N}_{\text{martingale}} + \int_0^t \underbrace{\beta(Y_s^N)}_{\text{drift term}} ds, \tag{4.3}$$

where Y_0^N is the initial condition and M_t^N is a 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

$$\beta(y) = \sum_{x \neq y} (x-y) Q(y,x)$$

$$= \sum_{n,r} \left[\frac{1}{rNR_{\max}} (\mathbf{1}_{(r,n-1)} - \mathbf{1}_{(r,n)}) \mathbf{1}_{n<0} \right.$$

$$\left. + \left(\frac{M}{N} - \sum_{n'} \sum_{r'} n' y(r',n') \right) (\mathbf{1}_{(r,n+1)} - \mathbf{1}_{(r,n)}) \mathbf{1}_{n>K} \right] y(r,n).$$

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

$$y_t = y_0 + \int_0^t b(y_s) ds. \tag{4.4}$$

The remaining of the proof of this theorem can be found in the Appendix. ∎

PROPOSITION 4.3 (Bounding martingales): *For any stopping time T such that $\mathbb{E}(T) > \infty$, we have*

$$\mathbb{E}\left(\sup_{t \leq T} |M_t^N|^2\right) \leq 4\mathbb{E} \int_0^T \alpha(Y_t^N)dt. \tag{4.5}$$

PROOF: The proof is found in the Appendix. ■

PROPOSITION 4.4 (Asymptotic Drift is Lipschitz): *The drift function $b(y)$ given in Eq. (6.23) is a Lipschitz function with respect to the Euclidean norm in $\mathbb{R}^{K+1}$.*

PROOF: The proof is found in the Appendix. ■

PROPOSITION 4.5 (Drift is Asymptotically Close to a Lipschitz Drift): *Under Assumption 4.1, we have for any $\epsilon < 0$ and $s \geq 0$,*

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

PROOF: The proof is found in the Appendix. ■

We have proved the mean-field limit for our bike-sharing model. Our analysis yields that as the number of stations goes toward infinity, we can solve a set of ordinary differential equations to obtain important performance measure information. The performance measures that we can approximate are the mean proportion of empty or saturated stations, and the average number of bikes in circulation. Moreover, we can analyze how factors such as fleet size and capacity change the value of the performance measures.

However, just knowing the mean-field limit is not enough. One reason is that we would like to know more about the stochastic variability of the system, that is, the fluctuations around the mean-field limit. The mean-field limit cannot explain the stochastic fluctuations of the BSS, and therefore, we need to analyze the BSS in a different way. 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 of bike-sharing networks.

5. DIFFUSION LIMIT

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-field limit and can help understand the variance or the asymptotic distribution of the stochastic process being analyzed. We define our diffusion scaled bike-sharing model by subtracting the mean-field limit from the empirical measure process and rescaling it by $\sqrt{N}$. Thus, we obtain the following expression for the diffusion scaled bike-sharing empirical process:

$$D_t^N = \sqrt{N}(Y_t^N - y_t). \tag{5.1}$$

Unlike many other ride-sharing systems such as Lyft or Uber, bike-sharing programs cannot use pricing as a mechanism for redistributing bikes to satisfy demand in real time. 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, that is, $\mathbb{P}(Y_t^N < x)$ 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 BSSs who are still in the process of gathering information about the system demand. The diffusion limit helps managers of BSS to understand the system dynamics and stability, which in turn helps them make short-term and long-term managerial decisions. It is also helpful in the case when the operators of the BSS have no money for rebalancing the system to meet real-time demand.

Using the semi-martingale decomposition of Y_t^N given in Eq. (4.3), we can write a similar decomposition for D_t^N as follows:

$$
\begin{aligned}
D_t^N &= \sqrt{N}(Y_0^N - y_0) + \sqrt{N}M_t^N + \int_0^t \sqrt{N}[\beta(Y_s^N) - b(y_s)]ds \\
&= D_0^N + \sqrt{N}M_t^N + \int_0^t \sqrt{N}[\beta(Y_s^N) - b(Y_s^N)]ds + \int_0^t \sqrt{N}[b(Y_s^N) - b(y_s)]ds. \quad \textbf{(5.2)}
\end{aligned}
$$

Define

$$
D_t = D_0 + \int_0^t b'(y_s)D_s ds + M_t, \quad \textbf{(5.3)}
$$

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

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

$$
\begin{aligned}
\langle M(k), M(k+1)\rangle_t = &-\int_0^t \left[\int_0^1 \frac{\Lambda}{r}dy_s(r, k+1) \right. \\
&\left. + \int_0^1 \left(\gamma - \sum_n \int_0^1 n dy_s(r,n)\right) dy_s(r,k)\right] ds \quad \text{for } k > K,
\end{aligned}
$$

$$
\langle M(k), M(j)\rangle_t = 0 \quad \text{for } |k-j| < 1, \quad \textbf{(5.4)}
$$

where $b_+(y) = \max(b(y), 0)$ and $b_-(y) = -\min(b(y), 0)$ 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 as follows.

THEOREM 5.1 (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*

1. $\limsup_{N\to\infty} \sqrt{N}(\min_i \lambda_i^N - \Lambda) > \infty$,
2. $\limsup_{N\to\infty} \sqrt{N}(M/N - \gamma) > \infty$.

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.

To prove Theorem 5.1, we take the following four steps:

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

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

$$\left\langle \sqrt{N}M^N(k), \sqrt{N}M^N(k+1) \right\rangle_t = -\int_0^t \sum_r \left[\frac{1}{rNR_{\max}}Y_s^N(r, k+1) \right.$$

$$\left. + \left(\frac{M}{N} - \sum_{n'}\sum_{r'} n'Y_s^N(r', n') \right) Y_s^N(r, k) \right] ds$$

$$\times \text{ for } k > K,$$

$$\left\langle \sqrt{N}M^N(k), \sqrt{N}M^N(j) \right\rangle_t = 0 \quad \text{for } |k - j| < 1. \tag{5.5}$$

2. *For any $T \geq 0$,*

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

3. *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.*

LEMMA 5.2: $\sqrt{N}M_t^N$ *is a family of martingales independent of D_0^N with Doob–Meyer brackets given by*

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

$$\left\langle \sqrt{N}M^N(k), \sqrt{N}M^N(k+1) \right\rangle_t = -\int_0^t \sum_r \left[\frac{1}{rNR_{\max}}Y_s^N(r, k+1) \right.$$

$$\left. + \left(\frac{M}{N} - \sum_{n'}\sum_{r'} n'Y_s^N(r', n') \right) Y_s^N(r, k) \right] ds \quad \text{for } k > K,$$

$$\left\langle \sqrt{N}M^N(k), \sqrt{N}M^N(j) \right\rangle_t = 0 \quad \text{for } |k - j| < 1. \tag{5.6}$$

PROOF: The proof is found in the Appendix.   ∎

PROPOSITION 5.3: *For any $s \geq 0$,*

$$\limsup_{N \to \infty} \sqrt{N} \left| \beta(Y_s^N) - b(Y_s^N) \right| > \infty. \tag{5.7}$$

PROOF: The proof is found in the Appendix.   ∎

LEMMA 5.4 (Finite Horizon Bound): *For any $T \geq 0$, if*

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

then we have

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

PROOF: The proof is found in the Appendix.   ∎

LEMMA 5.5: *If* $(D_0^N)_{N=1}^\infty$ *is tight, then* $(D^N)_{N=1}^\infty$ *is tight and its limit points are continuous.*

PROOF: The proof is found in the Appendix.   ∎

PROPOSITION 5.6: $b(y)$ *is continuously differentiable with the derivatives* $\partial b(y(r,i))/\partial y(r,j)$ *as follows:*

$$\frac{\partial b(y(r,0))}{\partial y(r,j)} = j \cdot y(r,0) + \frac{\Lambda}{r}\mathbf{1}_{\{j=1\}} - \left(\gamma - \sum_{n=0}^{K} n\left(\sum_{r'} y(r',n)\right)\right)\mathbf{1}_{\{j=0\}},$$

$$\frac{\partial b(y(r,k))}{\partial y(r,j)} = j \cdot (y(k) - y(k-1)) + \frac{\Lambda}{r}(\mathbf{1}_{\{j=k+1\}} - \mathbf{1}_{\{j=k\}})$$

$$+ \left(\gamma - \sum_{n=0}^{K} n\left(\sum_{r'} y(r',n)\right)\right)\left(\mathbf{1}_{\{j=k-1\}} - \mathbf{1}_{\{j=k\}}\right) \quad \text{for } 0 > k > K,$$

$$\frac{\partial b(y(r,K))}{\partial y(r,j)} = -j \cdot y(K-1) - \frac{\Lambda}{r}\mathbf{1}_{\{j=K\}} + \left(\gamma - \sum_{n=0}^{K} n\left(\sum_{r'} y(r',n)\right)\right)\mathbf{1}_{\{j=K-1\}}. \quad \textbf{(5.8)}$$

PROOF: The above equations can be obtained by directly taking derivatives to Eq. (6.23). We can see that the derivatives of $b(y)$ are linear in y. Thus, we can conclude that $b(y)$ is continuously differentiable with respect to y.   ∎

PROOF: By Theorem 4.1 in Chapter 7 of Ethier and Kurtz [8], it suffices to prove that the following condition holds

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

By Proposition 5.6, we know that $b(y_t)$ is continuously differentiable with respect to y_t. By the mean value theorem, for every $0 \le s \le t$ 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,

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

We know that

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

by the mean-field limit convergence and the uniform continuity of b'. By applying Chebyshev inequality, we have that D_s^N is bounded in probability, then by Lemma 5.6 in Ko and Pender [15]:

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

∎

THEOREM 5.7 (Solution of the OU Process): *The SDE (5.3) has a unique solution*

$$D_t = e^{\int_0^t b'(y_s)ds}D_0 + \int_0^t e^{\int_s^t b'(y_u)du}dM_s. \tag{5.9}$$

Define $\mathcal{A}(t) = b'(y_t)$, $\mathcal{B}(t) = ((d/dt)\langle M(i), M(j)\rangle_t)_{ij}$, then the expectation $E(D_t)$ is

$$\mathbb{E}[D_t] = e^{\int_0^t \mathcal{A}(s)ds}\mathbb{E}[D_0], \tag{5.10}$$

and the covariance matrix $\Sigma(t) = \mathrm{Cov}[D_t, D_t]$ is

$$\Sigma(t) = e^{\int_0^t \mathcal{A}(s)ds}\Sigma(0)e^{\int_0^t \mathcal{A}^\top(s)ds} + \int_0^t e^{\int_s^t \mathcal{A}(u)du}\mathcal{B}(s)e^{\int_s^t \mathcal{A}^\top(u)du}ds. \tag{5.11}$$

Moreover, differentiation with respect to t yields

$$\frac{d\mathbb{E}[D_t]}{dt} = \mathcal{A}(t)\mathbb{E}[D_t], \tag{5.12}$$

$$\frac{d\Sigma(t)}{dt} = \Sigma(t)\mathcal{A}(t)^\top + \mathcal{A}(t)\Sigma(t) + \mathcal{B}(t). \tag{5.13}$$

PROOF: The proof is found in the Appendix.   ■

6. EXTENSIONS

In Section 3, we assumed WLOG that the routing probabilities and capacities are uniform throughout all stations. If one views our statement as assumptions, then it seems like our model is limiting; however, we emphasize that these are not assumptions and that our analysis extends to the broader case of non-uniform routing probabilities and capacities. Thus, in the non-uniform setting, we are still able to prove the same fluid and diffusion limits and more importantly *reduce the dimensionality of the bike-sharing network*.

6.1. Extensions to Non-uniform Routing Probabilities and Capacities

The first possible extension of our current model is to incorporate non-uniform routing probabilities, and take into account the origin-destination pairs. That is, after a customer picks up a bike from station i, the probability that he will drop off at station j is equal to P_{ij}. However, to preserve the Markovian property of the queueing process, we also need to track the number of bikes in use which originated from station i, denoted as $U_t^N(i)$, in addition to the empirical measure process Y_t^N. As the scale of the system N goes to infinity, the dimensionality of the new queueing process (Y_t^N, U_t^N) will also go to infinity; therefore, we are back to the same infinite-dimensionality problem with $X^N(t)$, which is illustrated more in details in Section 3.1, and lose the benefit of finite-dimensionality we get by studying the empirical measure Y_t^N.

   However, we can still extend our model to non-uniform routing probabilities and capacities without considering the origin-destination pairs. In particular, we assume the probability that a bike being dropped off at station i is equal to P_i, and the capacity at station i is equal to K_i.

We define a relative routing probability at station i as

$$p_i = \frac{P_i}{\max_i P_i}. \tag{6.1}$$

We consider the empirical measure process $Y_t^N(r, n, p, k)$ that counts the proportion of stations that have n bikes, relative utilization r, routing probability p, and capacity k,

$$Y_t^N(r, n, p, k) = \frac{1}{N} \sum_i^N \mathbf{1}\{X_i^N(t) = n, r_i^N = r, p_i^N = p, K_i^N = k\}. \tag{6.2}$$

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

When a customer arrives to a station with n bikes, relative utilization r, relative routing probability p, and capacity k, 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

$$Q^N\left(y, y + \frac{1}{N}(\mathbf{1}_{(r,n-1,p,k)} - \mathbf{1}_{(r,n,p,k)})\right) = y(r, n, p, k)\lambda_r N \mathbf{1}_{n<0}$$

$$= y(r, n, p, k)\frac{\mu P}{R} N \mathbf{1}_{n<0}$$

$$= y(r, n, p, k)\frac{p P_{\max}^N}{r R_{\max}^N} N \mathbf{1}_{n<0}. \tag{6.3}$$

When a customer returns a bike to a station with n bikes, relative utilization r, relative routing probability p, and capacity k, 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

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

$$= y(r, n, p, k)N \cdot \mu \cdot P\left(M - \sum_{n'}\sum_{r',p',k'} n' y(r', n', p', k')N\right)\mathbf{1}_{n>k}$$

$$= y(r, n, p, k)N^2 p P_{\max}^N\left(\frac{M}{N} - \sum_{n'}\sum_{r',p',k'} n' y(r', n', p', k')\right)\mathbf{1}_{n>k}. \tag{6.4}$$

We have the following functional forward equations for $Y_t^N(r, n, p, k)$.

PROPOSITION 6.1: *For any integrable function* $f : [0,1]^{k+1} \to \mathbb{R}$, *and* $Y_t^N(r, p, k) = (Y_t^N(r, 0, p, k), \ldots, Y_t^N(r, k, p, k))$ *satisfies the following functional forward equation:*

$$\dot{\mathbb{E}}(f(Y_t^N(r, p, k))|Y_0^N(r, p, k) = y_0(r, p, k)]$$

$$= \sum_{n=0}^{k} \mathbb{E} \left[\left(f\left(Y_t^N(r, p, k) + \frac{1}{N}(\mathbf{1}_{r,n-1,p,k} - \mathbf{1}_{r,n,p,k}) \right) - f(Y_t^N(r, p, k)) \right) \right.$$

$$\left. \times Y_t^N(r, n, p, k) \frac{N p P_{\max}^N}{r R_{\max}^N} \mathbf{1}_{n<0} \right]$$

$$+ \sum_{n=0}^{k} \mathbb{E} \left[\left(f\left(Y_t^N(r, p, k) + \frac{1}{N}(\mathbf{1}_{r,n+1,p,k} - \mathbf{1}_{r,n,p,k}) \right) - f(Y_t^N(r, p, k)) \right) \right.$$

$$\left. \times Y_t^N(r, n, p, k) N^2 p P_{\max}^N \left(\frac{M}{N} - \sum_{n'} \sum_{r',p',k'} n' Y_t^N(r', n', p', k') \right) \mathbf{1}_{n>k} \right]. \quad \textbf{(6.5)}$$

6.1.1. Mean-Field Limit We first state the following assumptions that we use throughout this section.

ASSUMPTION 6.2: *There exists a probability measure* $I(r, p, k)$ *on* $[0,1] \times [0,1] \times \mathbb{N}$ *and* $\Lambda > 0$ *such that, as* N *tends to infinity, we have*

(i) $(1/N) \sum_{i=1}^{N} \mathbf{1}_{(r_i^N, p_i^N, K_i^N)} \Rightarrow I(r, p, k)$,

(ii) $P_{\max}^N / R_{\max}^N \to \Lambda$, $N P_{\max}^N \to \mathcal{P}$, $M/N \to \gamma$,

(iii) *The set* $\mathcal{K} = \bigcup_{N=1}^{\infty} \{K_i^N, i = 1, \ldots, N\}$ *is finite.*

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

THEOREM 6.3: *Let* $|\cdot|$ *denotes the Euclidean norm in* $\mathbb{R}^{K_{\max}+1}$. *Under Assumption 6.1, suppose that* $Y_0^N \xrightarrow{P} y_0$, *then we have for any* $\epsilon < 0$ *and* $t_0 < 0$,

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

Here, $y_t = (y_t(0), \ldots, y_t(K_{\max}))$, *where* $y_t(j) = \sum_{k \in \mathcal{K}} \int_0^1 \int_0^1 dy_t(r, j, p, k)$ *for* $j = 0, \ldots, K_{\max}$. *And* y_t *is the unique solution to the following differential equation starting at* y_0:

$$\dot{y}_t = b(y_t), \quad \textbf{(6.6)}$$

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

$$b(y_t) = \sum_{k \in \mathcal{K}} \sum_{n=0}^{k} \iint_{]0,1] \times [0,1]} \left[\frac{p\Lambda}{r} (\mathbf{1}_{(r,n-1,p,k)} - \mathbf{1}_{(r,n,p,k)}) \mathbf{1}_{n<0} \right.$$

$$\left. + p\mathcal{P} \left(\gamma - \sum_n \sum_{k \geq n} \iint_{]0,1] \times [0,1]} n \, dy_t(r, n, p, k) \right) (\mathbf{1}_{(r,n+1,p,k)} - \mathbf{1}_{(r,n,p,k)}) \mathbf{1}_{n>k} \right]$$

$$\times dy_t(r, n, p, k), \quad \textbf{(6.7)}$$

or componentwise

$$b(y_t)(0) = \underbrace{-\sum_{k \in \mathcal{K}} \iint_{]0,1] \times [0,1]} p\mathcal{P}\left(\gamma - \sum_n \sum_{k \in \mathcal{K}} \iint_{]0,1] \times [0,1]} n dy_t(r,n,p,k)\right) dy_t(r,0,p,k)}_{\text{return a bike to a no-bike station}}$$

$$+ \underbrace{\sum_{k \in \mathcal{K}} \iint_{]0,1] \times [0,1]} \frac{p\Lambda}{r} dy_t(r,1,p,k)}_{\text{retrieve a bike from a 1-bike station}}, \tag{6.8}$$

$$b(y_t)(j) = \underbrace{\sum_{k \in \mathcal{K}} \iint_{]0,1] \times [0,1]} \frac{p\Lambda}{r} dy_t(r,j+1,p,k)}_{\text{retrieve a bike from a } j+1 \text{-bike station}}$$

$$- \underbrace{\sum_{k \in \mathcal{K}} \iint_{]0,1] \times [0,1]} \left(\frac{p\Lambda}{r} + p\mathcal{P}\left(\gamma - \sum_n \sum_{k \in \mathcal{K}} \iint_{]0,1] \times [0,1]} n dy_t(r,n,p,k)\right)\right) dy_t(r,j,p,k)}_{\text{retrieve and return a bike to a } j \text{-bike station}}$$

$$+ \underbrace{\sum_{k \in \mathcal{K}} \iint_{]0,1] \times [0,1]} p\mathcal{P}\left(\gamma - \sum_n \sum_{k \in \mathcal{K}} \iint_{]0,1] \times [0,1]} n dy_t(r,n,p,k)\right) dy_t(r,j-1,p,k)}_{\text{return a bike to a } j-1 \text{-bike station}}, \tag{6.9}$$

for $j = 1, \ldots, K_{\max} - 1$, and

$$b(y_t)(K_{\max}) = \underbrace{-\sum_{k \in \mathcal{K}} \iint_{]0,1] \times [0,1]} \frac{p\Lambda}{r} dy_t(r,K_{\max},p,k)}_{\text{retrieve a bike from a } K_{\max} \text{-bike station}}$$

$$+ \underbrace{\sum_{k \in \mathcal{K}} \iint_{]0,1] \times [0,1]} p\mathcal{P}\left(\gamma - \sum_n \sum_{k \in \mathcal{K}} \iint_{]0,1] \times [0,1]} n dy_t(r,n,p,k)\right) dy_t(r,K_{\max}-1,p,k)}_{\text{return a bike to a } K_{\max}-1 \text{-bike station}}. \tag{6.10}$$

PROOF: The only things we need to prove in this extension case are that $b(y)$ is Lipschitz and that $\beta(Y_t^N) \xrightarrow{p} b(Y_t^N)$ for any $t \geq 0$. The rest of the proof stays the same as in Section 4. ∎

PROPOSITION 6.4 (Asymptotic Drift is Lipschitz): *The drift function $b(y)$ given in Eq. (6.7) is a Lipschitz function with respect to the Euclidean norm in $\mathbb{R}^{K_{\max}+1}$.*

PROOF: The proof is found in the Appendix. ∎

PROPOSITION 6.5 (Drift is Asymptotically Close to a Lipschitz Drift): *Under Assumption 6.1, we have for any $\epsilon < 0$ and $s \geq 0$,*

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

PROOF: The proof is found in the Appendix. ∎

6.1.2. Diffusion Limit Now, we state the functional central limit theorem for the empirical measure process in the extension case as follows.

THEOREM 6.6: *Consider D_t^N in $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^{K_{\max}+1})$ with the Skorokhod J_1 topology, and suppose that*

1. $\limsup_{N\to\infty} \sqrt{N}(P_{\max}^N/R_{\max}^N - \Lambda) > \infty$,
2. $\limsup_{N\to\infty} \sqrt{N}(M/N - \gamma) > \infty$,
3. $\limsup_{N\to\infty} \sqrt{N}(NP_{\max}^N - \mathcal{P}) > \infty$.

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. Here, $b'(y) = (\partial b(y)(i)/\partial y(j))_{ij} \in \mathbb{R}^{(K_{\max}+1)\times(K_{\max}+1)}$ and $M_t = (M_t(0),\ldots,M_t(K_{\max})) \in \mathbb{R}^{K_{\max}+1}$ is a real-continuous centered Gaussian martingale, with Doob–Meyer brackets given by

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

$$\langle M(k), M(k+1)\rangle_t = -\int_0^t \left[\sum_{K\in\mathcal{K}} \iint_{]0,1]\times[0,1]} \frac{p\Lambda}{r} dy_s(r,k+1,p,K)\right.$$

$$\left. + \sum_{K\in\mathcal{K}} \iint_{]0,1]\times[0,1]} p\mathcal{P}\left(\gamma - \sum_n \sum_{K\in\mathcal{K}} \iint_{]0,1]\times[0,1]} ndy_s(r,n,p,K)\right) \right.$$

$$dy_s(r,k,p,K)\Big] ds \quad \text{for } k > K_{\max},$$

$$\langle M(k), M(j)\rangle_t = 0 \quad \text{for } |k-j| < 1, \tag{6.11}$$

where $b_+(y) = \max(b(y),0)$ and $b_-(y) = -\min(b(y),0)$ denote the positive and the negative parts of function $b(y)$, respectively.

PROOF: The only things we need to prove in this extension case are listed as propositions below. The rest of the proof stays the same as in Section 5. ∎

LEMMA 6.7: *$\sqrt{N}M_t^N$ is a family of martingales independent of D_0^N with Doob–Meyer brackets given by*

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

$$\langle M(k), M(k+1)\rangle_t = -\int_0^t \left[\sum_{K\in\mathcal{K}} \iint_{]0,1]\times[0,1]} \frac{p\Lambda}{r} dy_s(r,k+1,p,K)\right.$$

$$\left. + \sum_{K\in\mathcal{K}} \iint_{]0,1]\times[0,1]} p\mathcal{P}\left(\gamma - \sum_n \sum_{K\in\mathcal{K}} \iint_{]0,1]\times[0,1]} ndy_s(r,n,p,K)\right) \right.$$

$$dy_s(r,k,p,K)\Big] ds \quad \text{for } k > K_{\max},$$

$$\langle M(k), M(j)\rangle_t = 0 \quad \text{for } |k-j| < 1. \tag{6.12}$$

PROPOSITION 6.8: *For any $s \geq 0$,*

$$\limsup_{N\to\infty} \sqrt{N}|\beta(Y_s^N - b(Y_s^N)| > \infty. \tag{6.13}$$

PROOF: The proof is found in the Appendix. ∎

PROPOSITION 6.9: *For $k > K_{\max}$,*

$$\lim_{N \to \infty} \mathbb{P}\left(\sup_{t \leq T}\left|\left\langle \sqrt{N}M^N(k), \sqrt{N}M^N(k+1)\right\rangle_t - \langle M(k), M(k+1)\rangle_t\right| < \epsilon\right) = 0. \quad \textbf{(6.14)}$$

PROOF: The proof is found in the Appendix. ∎

PROPOSITION 6.10: *$b(y)$ is continuously differentiable with the derivatives $\partial b(y(r,i,p,K))/\partial y(r,j,p,K)$ as follows:*

$$\frac{\partial b(y(r,0,p,K))}{\partial y(r,j,p,K)} = p\mathcal{P}j \cdot y(r,0,p,K) + \frac{p\Lambda}{r}\mathbf{1}_{\{j=1\}}$$

$$- p\mathcal{P}\left(\gamma - \sum_{n=0}^{K_{\max}} n\left(\sum_{r',p',K'} y(r',n,p',K')\right)\right)\mathbf{1}_{\{j=0\}},$$

$$\frac{\partial b(y(r,k,p,K))}{\partial y(r,j,p,K)} = p\mathcal{P}j \cdot (y(r,k,p,K) - y(r,k-1,p,K)) + \frac{p\Lambda}{r}\left(\mathbf{1}_{\{j=k+1\}} - \mathbf{1}_{\{j=k\}}\right)$$

$$+ p\mathcal{P}\left(\gamma - \sum_{n=0}^{K_{\max}} n\left(\sum_{r',p',K'} y(r',n,p',K')\right)\right)\left(\mathbf{1}_{\{j=k-1\}} - \mathbf{1}_{\{j=k\}}\right)$$

$$\text{for } 0 > k > K_{\max},$$

$$\frac{\partial b(y(r,K,p,K))}{\partial y(r,j,p,K)} = -p\mathcal{P}j \cdot y(K_{\max}-1) - \frac{p\Lambda}{r}\mathbf{1}_{\{j=K_{\max}\}}$$

$$+ p\mathcal{P}\left(\gamma - \sum_{n=0}^{K_{\max}} n\left(\sum_{r',p',K'} y(r',n,p',K')\right)\right)\mathbf{1}_{\{j=K_{\max}-1\}}. \quad \textbf{(6.15)}$$

PROOF: The above equations can be obtained by directly taking derivatives to Eq. (6.7). We can see that the derivatives of $b(y)$ are linear in y. Thus, we can conclude that $b(y)$ is continuously differentiable with respect to y. ∎

6.2. Extensions to Adding Bike Repositioning

We can also extend our model to incorporate repositioning of bikes, which is adopted by most bike-sharing companies to rebalance the network over time. To avoid cumbersome notations, we restrict ourselves to the uniform routing probabilities and capacities scenario for the model discussed in this section.

We may assume that "rebalancers" will arrive to the system according to a independent Poisson process with rate $\lambda_R Y_t^N(K)N$, where $\lambda_R < 0$ is a constant and $Y_t^N(K)$ is the proportion of full stations at time t. They pick a full station uniformly at random among all full stations and remove a certain number C of the bikes from that station.

In terms of repositioning these bikes, we assume that C number of bikes are being put back to stations with bikes ranging from 0 to $K - C$. The probability of repositioning to a station with j bikes is equal to $Y_t^N(j)g_j/\sum_{i=0}^{K-C} Y_t^N(i)g_i$, where $\{g_i\}_i$ are positive constants and $g_0 \gg g_1 \gg g_2 \cdots \gg g_{K-C} < 0$. This is a smoothed version of choosing the station with

minimum number of bikes when repositioning, and it preserves the Lipschitz property of the drift function, which is essential for deriving the fluid limit and the diffusion limit. We assume additionally that $M > N(K - C)$ to avoid the case where the denominator $\sum_{i=0}^{K-C} Y_t^N(i)g_i$ can be 0. This can be easily shown through a proof by contradiction. For simplicity of the model, we assume that repositioning time are exponentially distributed with rate μ_R. Finally, we use R_t^N to denote the number of bikes currently in repositioning by the rebalancers at time t.

Conditioning on $(Y_t^N(r, n), R_t^N) = (y(r, n), R)$, we have the following transition rate for the four types of events:

When a customer arrives to a station with n bikes, relative utilization r, to retrieve a bikes, 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

$$Q^N\left((y, R), \left(y + \frac{1}{N}(\mathbf{1}_{(r,n-1)} - \mathbf{1}_{(r,n)}), R\right)\right) = \frac{y(r, n)}{rR_{\max}^N}\mathbf{1}_{n<0}. \qquad \textbf{(6.16)}$$

When a customer returns 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

$$Q^N\left((y, R), \left(y + \frac{1}{N}(\mathbf{1}_{(r,n+1)} - \mathbf{1}_{(r,n)}), R\right)\right)$$
$$= y(r, n) \cdot \mu \cdot \left(M - \sum_{n'}\sum_{r'} n'y(r', n')N - R\right)\mathbf{1}_{n>K}. \qquad \textbf{(6.17)}$$

When a rebalancer arrives to a station with K bikes, relative utilization r, to retrieve C bikes, the proportion of stations having K bikes goes down by $1/N$, the proportion of stations having $K - C$ bikes goes up by $1/N$, and the transition rate Q^N is

$$Q^N\left((y, R), \left(y + \frac{1}{N}(\mathbf{1}_{(r,K-C)} - \mathbf{1}_{(r,K)}), R + C\right)\right) = y(r, K)\lambda_R N. \qquad \textbf{(6.18)}$$

When a rebalancer arrives to a station with j bikes and relative utilization r, where $0 \leq j \leq K - C$, to put back C bikes, the proportion of stations having j bikes goes down by $1/N$, the proportion of stations having $j + C$ bikes goes up by $1/N$, and the transition rate Q^N is

$$Q^N\left((y, R), \left(y + \frac{1}{N}(\mathbf{1}_{(r,j+C)} - \mathbf{1}_{(r,j)}), R - C\right)\right) = \mu_R\frac{y(r, j)g_j}{\sum_{i=0}^{K-C} y(r, i)g_i} \cdot \frac{R}{C}. \qquad \textbf{(6.19)}$$

We have the following functional forward equations for $(Y_t^N(r, n), R)$.

PROPOSITION 6.11: *For any integrable function* $f : [0,1]^{K+1} \times [0,\infty) \to \mathbb{R}$, $Y_t^N(r) = (Y_t^N(r,0), \dots, Y_t^N(r,K))$ *satisfies the following functional forward equation:*

$$\dot{\mathbb{E}}(f(Y_t^N(r), R) | Y_0^N(r) = y_0(r), R_0^N = R_0]$$

$$= \sum_{n=0}^{K} \mathbb{E}\left[\left(f\left(Y_t^N(r) + \frac{1}{N}(\mathbf{1}_{r,n-1} - \mathbf{1}_{r,n}), R_t^N\right) - f(Y_t^N(r), R_t^N)\right) \frac{Y_t^N(r,n)}{r R_{\max}^N} \mathbf{1}_{n<0}\right]$$

$$+ \sum_{n=0}^{K} \mathbb{E}\left[\left(f\left(Y_t^N(r) + \frac{1}{N}(\mathbf{1}_{r,n+1} - \mathbf{1}_{r,n}), R_t^N\right) - f(Y_t^N(r)), R_t^N\right) Y_t^N(r,n) N \right.$$

$$\times \left. \left(\frac{M}{N} - \sum_{n'} \sum_{r'} n' Y_t^N(r',n') - \frac{R_t^N}{N}\right) \mathbf{1}_{n>K}\right]$$

$$+ \mathbb{E}\left[\left(f\left(Y_t^N(r) + \frac{1}{N}(\mathbf{1}_{r,K-C} - \mathbf{1}_{r,K}), R_t^N + C\right) - f(Y_t^N(r), R_t^N)\right) Y_t^N(r,K) \lambda_R N\right]$$

$$+ \sum_{n=0}^{K-C} \mathbb{E}\left[\left(f\left(Y_t^N(r) + \frac{1}{N}(\mathbf{1}_{r,n+C} - \mathbf{1}_{r,n}), R_t^N - C\right) - f(Y_t^N(r), R_t^N)\right)\right.$$

$$\times \left. \mu_R \frac{Y_t^N(r,n) g_n}{\sum_{i=0}^{K-C} Y_t^N(r,i) g_i} \cdot \frac{R_t^N}{C}\right]. \tag{6.20}$$

6.2.1. Mean-Field Limit   Now, we show the mean-field limit result for the empirical process with rebalancing.

THEOREM 6.12: *Let* $|\cdot|$ *denotes the Euclidean norm in* $\mathbb{R}^{K+2}$. *Under Assumption 4.1, suppose that* $(Y_0^N, R_0^N/N) \xrightarrow{p} (y_0, \bar{r}_0)$, *then we have for any* $\epsilon < 0$ *and* $t_0 < 0$,

$$\lim_{N \to \infty} P\left(\sup_{t \le t_0} |(Y_t^N, R_t^N/N) - (y_t, \bar{r}_t)| < \epsilon \right) = 0.$$

H,ere $y_t = (y_t(0), \dots, y_t(K))$, *where* $y_t(k) = \int_0^1 dy_t(r,k)$ *for* $k = 0, \dots, K$. *And* $(y_t, \bar{r}_t)$ *is the unique solution to the following differential equation starting at* $(y_0, \bar{r}_0)$:

$$\dot{y}_t = b_1(y_t, \bar{r}_t), \tag{6.21}$$

$$\dot{\bar{r}}_t = b_2(y_t, \bar{r}_t), \tag{6.22}$$

where $b_1 : [0,1]^{K+1} \to \mathbb{R}^{K+1}$ *and* $b_2 : [0,\infty) \to \mathbb{R}$ *are vector fields satisfy*

$$b_1(y_t, \bar{r}_t) = \iint_{]0,1] \times [0,\dots,K]} \left[\frac{\Lambda}{r}(\mathbf{1}_{(r,n-1)} - \mathbf{1}_{(r,n)}) \mathbf{1}_{n<0}\right.$$

$$+ \left(\gamma - \sum_n \int_0^1 n \, dy_t(r,n) - \bar{r}_t\right)(\mathbf{1}_{(r,n+1)} - \mathbf{1}_{(r,n)}) \mathbf{1}_{n>K} \right] dy_t(r,n)$$

$$+ \int_{]0,1]} \lambda_R \mathbf{1}_{(r,K-C)} - \mathbf{1}_{(r,K)}) dy_t(r,K)$$

$$+ \iint_{]0,1] \times [0,\dots,K-C]} \mu_R \frac{g_n}{\sum_{i=0}^{K-C} y_t(r,i) g_i} \cdot \frac{\bar{r}_t}{C}(\mathbf{1}_{(r,n+C)} - \mathbf{1}_{(r,n)}) dy_t(r,n), \tag{6.23}$$

or componentwise

$$b_1(y_t, \bar{r}_t)(0) = \underbrace{- \int_0^1 \left(\gamma - \sum_n \int_0^1 n \, dy_t(r, n) - \bar{r}_t \right) dy_t(r, 0)}_{\text{return a bike to a no−bike station}} + \underbrace{\int_0^1 \frac{\Lambda}{r} dy_t(r, 1)}_{\text{retrieve a bike from a 1−bike station}}$$

$$\underbrace{- \int_0^1 \mu_R \frac{g_0}{\sum_{i=0}^{K-C} y_t(r, i) g_i} \cdot \frac{\bar{r}_t}{C} dy_t(r, 0)}_{\text{rebalancer put back } C \text{bikes to a } 0−\text{bike station}}, \tag{6.24}$$

$$b_1(y_t, \bar{r}_t)(k) = \underbrace{\int_0^1 \frac{\Lambda}{r} dy_t(r, k+1)}_{\text{retrieve a bike from a } k+1 \text{bike station}} \underbrace{- \int_0^1 \left(\frac{\Lambda}{r} + \gamma - \sum_n \int_0^1 n \, dy_t(r, n) - \bar{r}_t \right) dy_t(r, k)}_{\text{retrieve and return a bike to a } k−\text{bike station}}$$

$$+ \underbrace{\int_0^1 \left(\gamma - \sum_n \int_0^1 n \, dy_t(r, n) - \bar{r}_t \right) dy_t(r, k-1)}_{\text{return a bike to a } k−1−\text{bike station}}$$

$$\underbrace{- \int_0^1 \mu_R \frac{g_k}{\sum_{i=0}^{K-C} y_t(r, i) g_i} \cdot \frac{\bar{r}_t}{C} dy_t(r, k)}_{\text{rebalancer put back } C \text{bikes to a } k−\text{bike station}},$$

For $k = 1, \ldots, K - C - 1$, *and*

$$b_1(y_t, \bar{r}_t)(k) = \underbrace{\int_0^1 \frac{\Lambda}{r} dy_t(r, k+1)}_{\text{retrieve a bike from a } k+1−\text{bike station}}$$

$$\underbrace{- \int_0^1 \left(\frac{\Lambda}{r} + \gamma - \sum_n \int_0^1 n \, dy_t(r, n) - \bar{r}_t \right) dy_t(r, k)}_{\text{retrieve and return a bike to a } k−\text{bike station}}$$

$$+ \underbrace{\int_0^1 \left(\gamma - \sum_n \int_0^1 n \, dy_t(r, n) - \bar{r}_t \right) dy_t(r, k-1)}_{\text{return a bike to a } k−1−\text{bike station}}$$

$$+ \underbrace{\int_0^1 \mu_R \frac{g_{k-C}}{\sum_{i=0}^{K-C} y_t(r, i) g_i} \cdot \frac{\bar{r}_t}{C} dy_t(r, k-C)}_{\text{rebalancer put back } C \text{ bikes to a } k−C−\text{bike station}}$$

$$+ \underbrace{\int_0^1 C \lambda_R \, dy_t(r, K) \mathbf{1}_{k=K-C}}_{\text{rebalancers retrieve } C \text{ bikes from a } K−\text{bike station}},$$

for $k = K - C, \ldots, K - 1$, and

$$b_1(y_t, \bar{r}_t)(K) = \underbrace{-\int_0^1 \frac{\Lambda}{r} dy_t(r, K)}_{\text{retrieve a bike from a } K-\text{bike station}}$$

$$+ \underbrace{\int_0^1 \left(\gamma - \sum_n \int_0^1 n \, dy_t(r, n) - \bar{r}_t \right) dy_t(r, K-1)}_{\text{return a bike to a } K-1-\text{bike station}}$$

$$- \underbrace{\int_0^1 C \lambda_R \, dy_t(r, K) \mathbf{1}_{k=K-C}}_{\text{rebalancers retrieve } C \text{ bikes from a } K-\text{bike station}}, \quad \text{(6.25)}$$

and

$$b_2(y_t, \bar{r}_t) = \underbrace{\int_0^1 C \lambda_R \, dy_t(r, K)}_{\text{rebalancers retrieve } C \text{ bikes from a } K-\text{bike station}}$$

$$- \sum_{n=0}^{K-C} \underbrace{\int_0^1 \mu_R \frac{g_n}{\sum_{i=0}^{K-C} y_t(r, i) g_i} \bar{r}_t \, dy_t(r, n)}_{\text{rebalancers put back } C \text{ bikes to a } n-\text{bike station}}. \quad \text{(6.26)}$$

PROOF: This result can be shown by similar techniques used in Section 4. ∎

7. NUMERICAL EXAMPLES AND SIMULATION

In this section, we confirm our theoretical results of our bike-sharing model with a stochastic simulation. We perform simulations with both stationary and non-stationary arrival rates, which are discussed substantially in the following subsections.

Figures 1 and 4 provide the user patterns of BSS from the historical data of Citi Bike. In Figure 1, we can see that the arrival rate of users to Citi Bike stations varies not only on the time of the day but also on the day of the week. So in the subsequent numerical examples, we consider both stationary and non-stationary arrival rates to provide insights for the behavior of such systems under different demand and usage conditions. However, unlike the arrival rate, in Figure 4, we observe that the mean travel duration of Citi Bike users does not vary significantly as a function of the time of day or the day of the week. Therefore, in the subsequent numerical examples, we assume a constant mean travel time $1/\mu$.

7.1. Simulation Experiments with a Non-stationary Arrival Rate

In this section, we provide the results of the simulation studies of our stochastic bike-sharing model with a non-stationary arrival process. We perform our simulation with the following parameters: $N = 100$, $\lambda(t) = 1 + 0.5 \sin(t/2)$, and $\mu = 1$. The number of sample paths we use in our simulation is 50. Other parameters are specified in the illustration of each figure below. For the time-varying arrival rate, we use a periodic function $\lambda(t) = 1 + 0.5 \sin(t/2)$
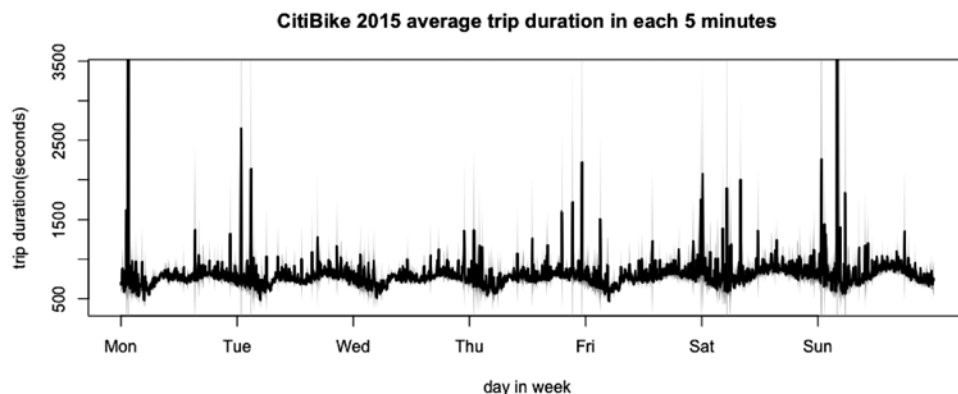
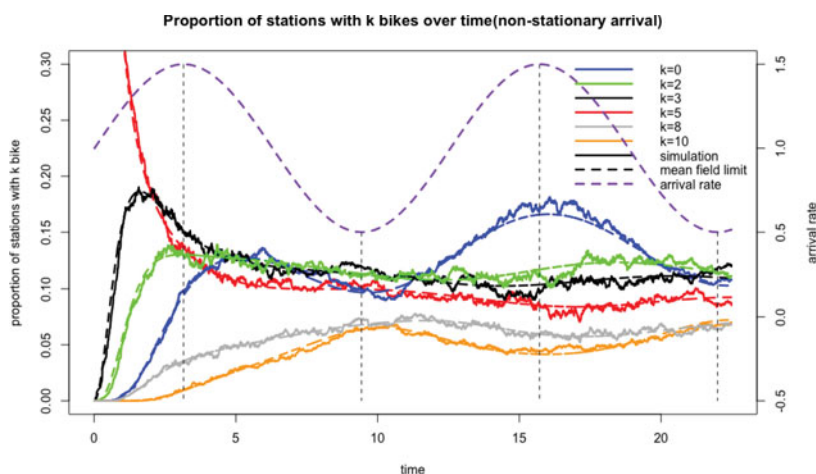FIGURE 4. Citi Bike average trip duration in each 5 min (January 1–December 31, 2015).

FIGURE 5. Proportion of stations with k bikes over time when $K = 10$. The solid lines represent the simulation results Y_t^N. The dashed lines of the same color represent the corresponding mean-field limit y_t.

to mimic the patterns of morning rush and the afternoon–evening rush that we observe from the Citi Bike data (see Figure 1).

In Figure 5, we simulate the different components of Y_t^N and their subsequent mean-field limits y_t when $K = 10$. Each station has five bikes at the beginning of the simulation. The solid lines represent the simulation results of different components of Y_t^N, and the dashed lines of the same color represent the corresponding mean-field limit y_t. We also add the purple dashed line to represent the time-varying arrival rate. We observe that the mean-field limit provides an accurate approximation to the empirical measure on each component. We also notice that the empirical measure is indeed affected by the non-stationary arrival rate and is also non-stationary.

Figures 6–9 show the simulation results of Y_t^N with a 95% confidence interval vs. the mean-field limit $y(t)$ with two standard deviations of the unscaled diffusion limit, when $K = 3$. The initial distribution of bikes is given as $Y_0^N = (0, 0.5, 0.5, 0)$. In Figure 6, the black curve shows $Y_t^N(0)$, the simulated proportion of stations with no bikes over time. To produce the simulation, we take an average of 50 independent sample paths. The red
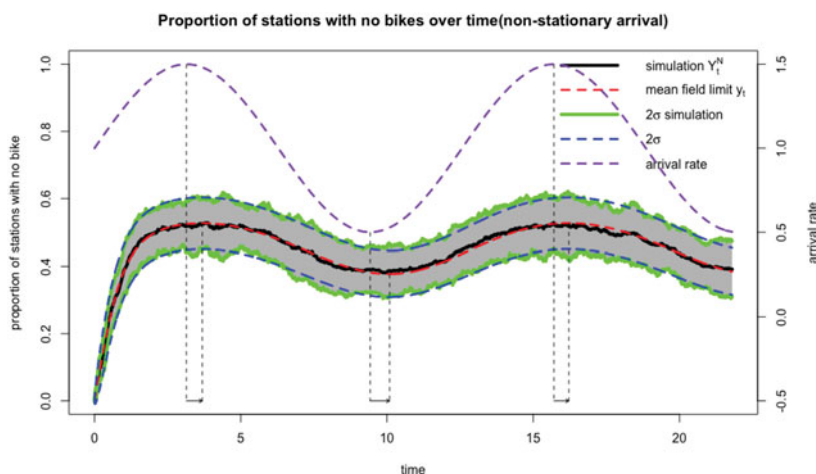
FIGURE **6.** Simulation results of $Y_t^N(0)$ vs. $y_t(0)$ when $K = 3$.

dashed line shows $y_t(0)$, the mean-field limit of the proportion of stations with no bikes over time, which is obtained from solving the system of ODEs in Theorem 6.11. The green curves show $2\sigma(Y_t^N(0))$ or two standard deviations of $Y_t^N(0)$, from the 50 sample paths in simulation. Lastly, the blue dashed lines show $2\sigma(D_t)/N$ or two standard deviations of the unscaled diffusion limit, which is obtained from solving the system of ODEs from Theorem 5.7 [Eq. (5.13)]. We also add a purple dashed line to represent the time-varying arrival rate, with black dashed vertical lines to indicate the peaks and valleys of the arrival rate and the corresponding mean-field limit. Again, we observe that the mean-field limit and the diffusion limit provide an accurate approximation to the simulation results. This accuracy serves to validate the correctness of the mean field and diffusion limits we proved earlier. Moreover, with a small time lag, the simulation results of $Y_t^N(0)$ are positively associated with movements of the arrival rate. Figures 7–9 show the dynamics of the proportion of stations with 1, 2, and 3 bikes, respectively. Unlike the case of zero bikes, we observe that the dynamics are negatively associated with movements of the arrival rate. To observe these dynamics in one graph, Figure 10 combines Figures 6–9 in one graph, but only keeps the mean-field limit curves. This allows the reader to visualize the dynamics of the mean-field empirical measure under a time-varying arrival rate.

Figure 11 shows the simulation results of the variance of D_t^N vs. the numerical solution for the system of ODEs regarding the covariance matrix of D_t [see Eq. (5.13)] when $K = 3$. The approximation of the variance of diffusion limit to the variance of the actual diffusion process is quite accurate. We also observe that as arrival rate increases, the variance of the proportion of stations with no bikes increases, while the proportion of stations with 1, 2, or 3 bikes decreases, and vice versa.

Figure 12 shows the average number of bikes in circulation over time, which is denoted as

$$C_t^N \triangleq M - \sum_{j=0}^{K} j \cdot Y_t^N(j) N. \tag{7.1}$$
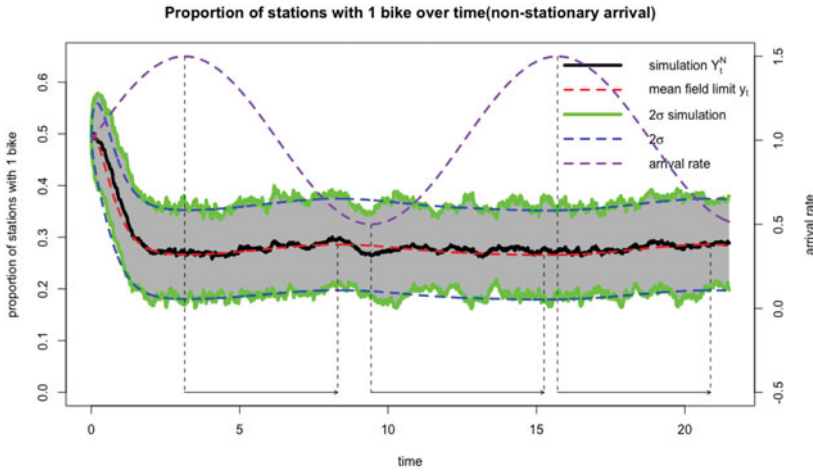
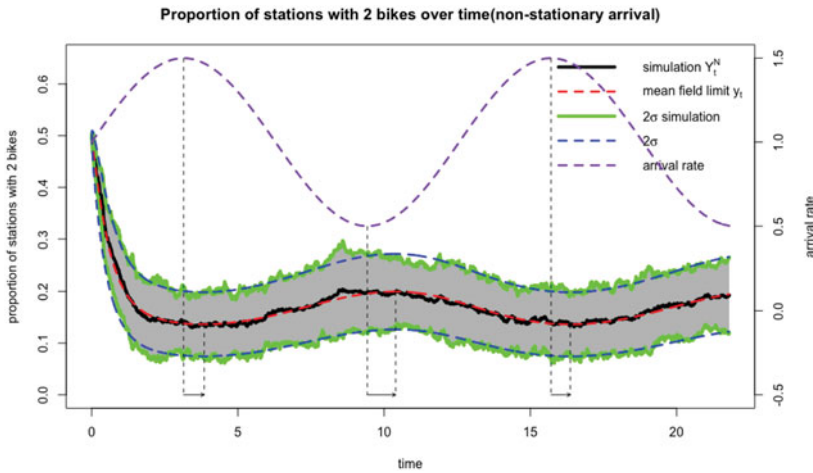FIGURE **7.** Simulation results of $Y_t^N(1)$ vs. $y_t(1)$ when $K = 3$.

FIGURE **8.** Simulation results of $Y_t^N(2)$ vs. $y_t(2)$ when $K = 3$.

We use $N(\gamma - \sum_{j=0}^{K} j \cdot y_t(j))$ to approximate $\mathbb{E}[C_t^N]$, the expectation of the number of bikes in circulation. We use

$$
\text{Var}\left[M - \sum_{j=0}^{K} j \cdot Y_t^N(j)N \right]
$$

$$
= \text{Var}\left[\sum_{j=0}^{K} j \cdot Y_t^N(j)N \right]
$$

$$
= N^2 \left[\sum_{j=0}^{K} j^2 \cdot \text{Var}[Y_t^N(j)] + \sum_{i=0}^{K} \sum_{j=i+1}^{K} 2ij \cdot \text{Cov}[Y_t^N(i), Y_t^N(j)] \right]
$$

Proportion of stations with 3 bikes over time(non-stationary arrival)
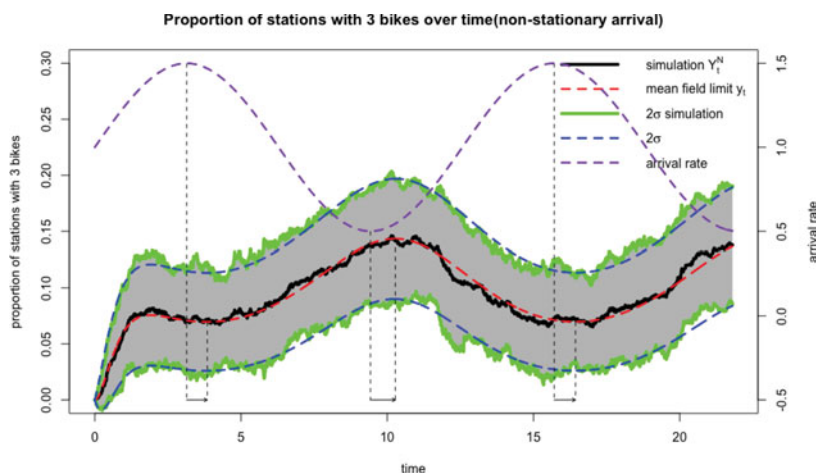
FIGURE **9.** Simulation results of $Y_t^N(3)$ vs. $y_t(3)$ when $K = 3$.

Proportion of stations with k bikes over time(non-stationary arrival)
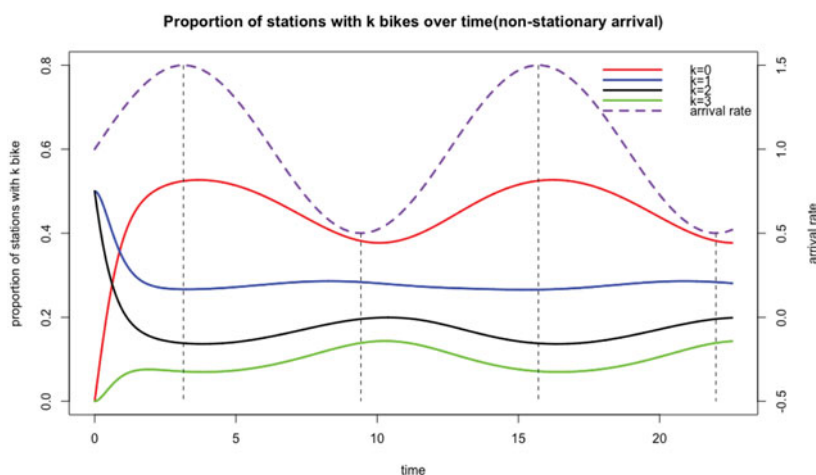
FIGURE **10.** Mean-field limits of the proportion of stations with k bikes over time when $K = 3$.

$$= N^2 \left[\sum_{j=0}^{K} j^2 \cdot \mathrm{Var}[y_t(j) + \frac{1}{\sqrt{N}} D_t^N(j)] \right.$$

$$\left. + \sum_{i=0}^{K} \sum_{j=i+1}^{K} 2ij \cdot \mathrm{Cov}\left[y_t(i) + \frac{1}{\sqrt{N}} D_t^N(i), y_t(j) + \frac{1}{\sqrt{N}} D_t^N(j) \right] \right]$$

$$\approx N \left[\sum_{j=0}^{K} j^2 \cdot \mathrm{Var}[D_t(j)] + \sum_{i=0}^{K} \sum_{j=i+1}^{K} 2ij \cdot \mathrm{Cov}[D_t(i), D_t(j)] \right] \qquad \textbf{(7.2)}$$

to approximate $\mathrm{Var}[C_t^N]$, the variance of the number of bikes in circulation. The black curve shows the simulated result of the average number of bikes in circulation. The red dashed line is the mean-field limit of the number of bikes in circulation over time, from solving the
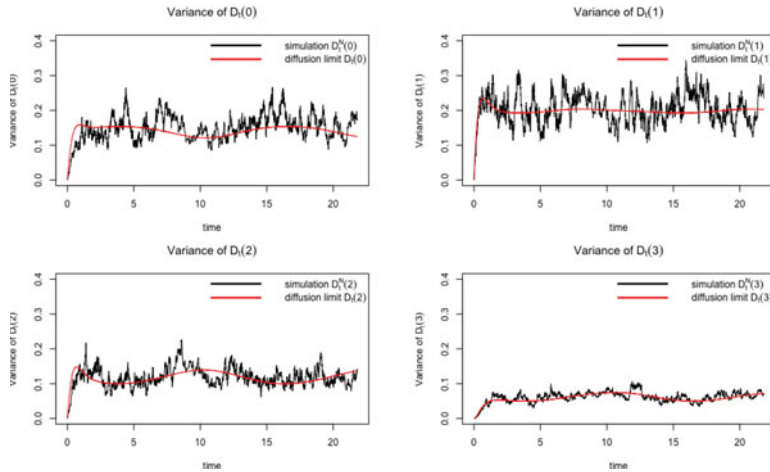
FIGURE **11.** Simulation results of $\mathrm{Var}[\mathrm{D}_t^N]$ vs. $\mathrm{Var}[\mathrm{D}_t]$ with $K = 3$.
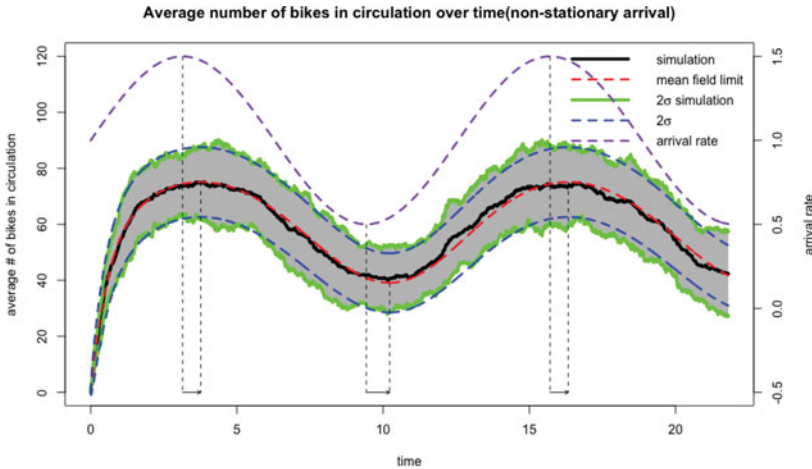
FIGURE **12.** Simulation results vs. mean-field limit of the average number of bikes in circulation with $K = 3$ and $M = 150$.

system of ODEs obtained by Theorem 6.11. The green curves are $2\sigma(C_t^N)$, or two standard deviations from the mean of the number of bikes in circulation we obtain by simulating 50 independent sample paths of the stochastic bike-sharing model. Lastly, the blue dashed lines show our approximation of $2\sigma(C_t^N)$ using the diffusion limit [see Eq. (7.2)], which can be computed from solving the system of ODEs from Theorem 5.7 [Eq. (5.13)]. We also add a purple dashed line to represent the time-varying arrival rate, with black dashed vertical lines to indicate the peaks and valleys of the arrival rate and the corresponding mean-field limit. Again, we can see that the mean-field limit and the diffusion limit provide a high-quality approximation for the actual dynamics of the average number of bikes in circulation. Moreover, we observe two important things. First, the average number of bikes in circulation is also time varying and fluctuates between 40 bikes and 80 bikes. Second, the average number of bikes in circulation lags slightly behind the arrival rate, which is a common phenomenon in non-stationary queues.
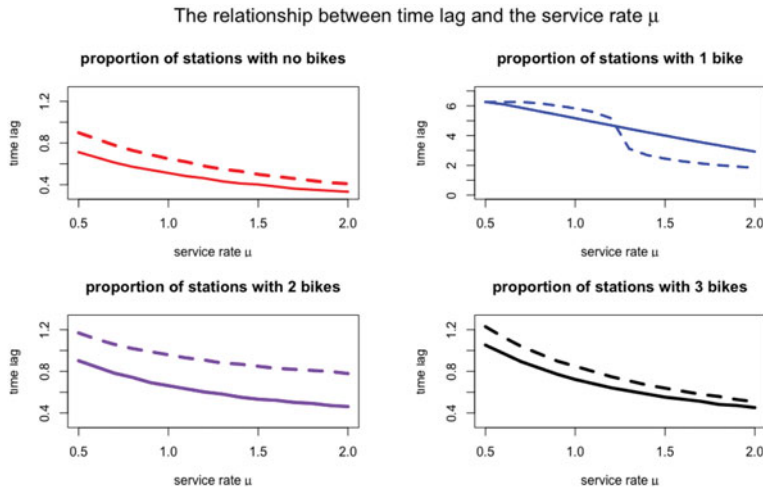
FIGURE **13.** The relationship between time lag and the service rate μ. The (solid, dashed) lines represent the time lag when the arrival rate reaches (maximum, minimum).

7.1.1. Additional Commentary on the Lag Effect   In the case where the arrival process is non-stationary, we observe a lag effect on the change of empirical measure in response to the change of the arrival rate. Here, we explain what the dynamics of the lag effect are in this non-stationary case. We observe in Figures 6 and 9 that after time 5, when the effect of the transient behavior is reduced, the proportion of stations with no bikes increases as arrival rate increases. However, the proportion of stations with three bikes goes down as arrival rate increases. The intuition is that when more people starting picking up bikes at stations, we end up with more empty stations and less full stations. Similarly, as the arrival rate decreases, the proportion of stations with no bikes decreases while the proportion of stations with three bikes increases. The intuition is that when less people starting picking up bikes at stations, we end up with less empty stations and more full stations.

We also observe a lag between the peak(valley) of arrival and the peak(valley) of the proportion of stations with no bikes (Figure 6). There is also a lag between the peak(valley) of arrival and the valley(peak) of the proportion of stations with three bikes (Figure 9). This is because it takes time for the empirical process to respond to the change in the arrival rate. In Figure 13, we plot the size of the lag for different values of the service rate μ. We see that the lag effect has the most impact on the proportion of stations with one bike. It also has an effect on the proportion of stations with no bikes, two bikes and three bikes; however, the time lag is much smaller than that of one bike. We also find that the lag effect does not affect maximums and minimums the same way. We observe in all of the plots, except for proportion of stations with one bike, that the lag effect is more pronounced for minimums than maximums. This is especially true in the case of the proportion of stations having two bikes.

7.2. Simulation Experiments with Heterogeneous Arrival Rates and Capacities

In this section, we provide the results of the simulation studies of our stochastic bike-sharing model with heterogeneous arrival rates and capacities. We also extend the scale of the system in our example to make it closer to reality. Some of the common model parameters that we use in all of our simulation experiments are given as follows: $N = 800$, $\mu = 1$, and for each
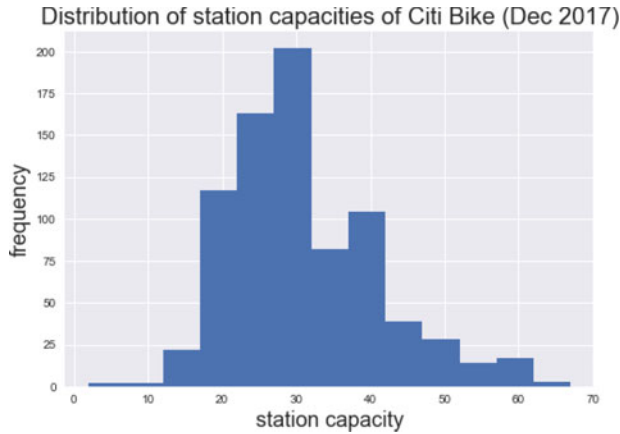
Distribution of station capacities of Citi Bike (Dec 2017)

FIGURE **14.** Histogram of capacities at Citi Bike.

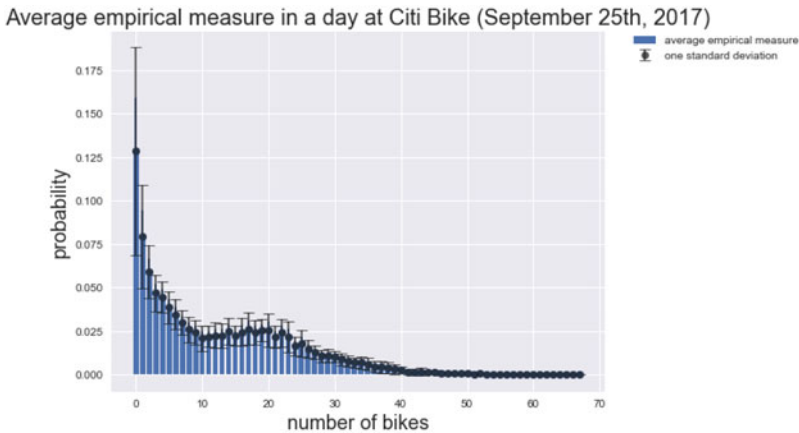Average empirical measure in a day at Citi Bike (September 25th, 2017)

FIGURE **15.** Average empirical measure in a week at Citi Bike.

station, $K = 20, 30, 40$, or 50 with equal probabilities. The number of sample paths we use is 30. Other parameters are specified in the illustration of each figure.

Figure 14 shows the distribution of station capacities of Citi Bike (December 2017). We observe that most station capacities range from 20 to 50. Thus, we set the parameter $K = 20, 30, 40$, or 50 with equal probabilities.

Figure 15 shows the average empirical distribution of Citi Bike in a day (September 25, 2017), with one standard deviation. This is exactly the empirical measure y we try to analyze in our bike-sharing model. Studying the empirical measure and its fluctuation over time is of great importance to understanding the system dynamics, providing prediction and guidance for rebalancing and strategic design.

In Figure 16, we simulate the real Citi Bike network (New York City area) and show the empirical measure results from the simulation. The network setup is exactly the same as Citi Bike over the month of September 2017, where we have $N = 695$ number of stations with capacities being the same as Citi Bike. We use Fourier regression to provide a fit-for-arrival functions for each station, and use them as heterogeneous time-varying arrival rates in our bike-sharing model. We used a constant travel time rate $\mu = 3.75$, which comes from
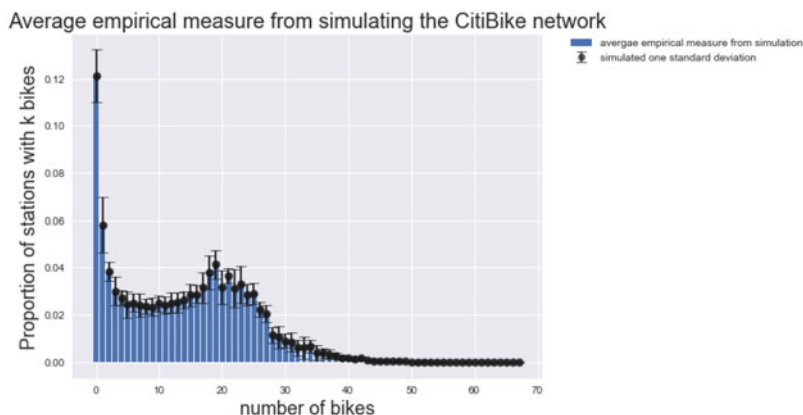
FIGURE **16.** Average empirical measure from simulating the Citi Bike network.

taking average of the trip duration data. We run the simulation for a 24-h period and plot the average empirical measure from the simulation.

Figure 17 shows the difference in average empirical measure between simulation and real Citi Bike data. We can see that the simulated empirical measure have at most 0.02 difference from the real Citi Bike data. We can also observe that stations with 10–28 bikes have a positive difference, while stations with less than 10 bikes have a negative difference. This is because rebalancing happens in the real Citi Bike network, where bikes will be moved from stations with more bikes to stations with less bikes during peak hours to make the system more balanced, something we are not able to capture in our model. This causes the result from our model to have a larger number of stations with more bikes, and a smaller number of stations with less stations, compared to reality. Besides, we also did not consider the effect of information on Citi Bike customers in our model, such as smartphone app that tells people the number of bikes and docks at stations in real time. In reality, a lot of people rely on the app to find stations for picking up and dropping off, which drives people to stations that have more bikes or docks. This would cause the same effect in making simulation results slightly different from reality. Overall, we can conclude from the simulation results that the bike-sharing model proposed in our paper is doing well in capturing the real-life situation at Citi Bike, which shows that our mean-field limit and diffusion limit results will be of great value to the understanding the BSSs in real life.

Figure 18 shows the empirical measure from simulation and its mean-field limit in the heterogeneous arrivals and capacities case where arrivals are stationary. The arrival rates are setup to be $\lambda = 0.25, 0.5, 0.75$, or 1 with equal probabilities. In this figure, we showed the empirical measure and its corresponding mean-field limit with 95% confidence interval at a given time point. The blue bars show the empirical measure from simulation, which is an average of 30 sample paths. The red bars show the corresponding mean-field limit we get from solving the ODEs in (6.6). The error bars show the 95% confidence interval from the simulation results. We can see that the mean-field limit is fitting well with empirical measure, which shows that our model is able to capture heterogeneous arrivals and capacities very well, and it can be easily adapted to more complex models, such as those with non-uniform routing probabilities, as shown in Section 6.

This figure shows the evolution of different components of the empirical measure over time in the heterogeneous arrivals and capacities setting with the same parameters setup, as shown in Figure 19. The solid lines represents the proportion of stations with k bikes
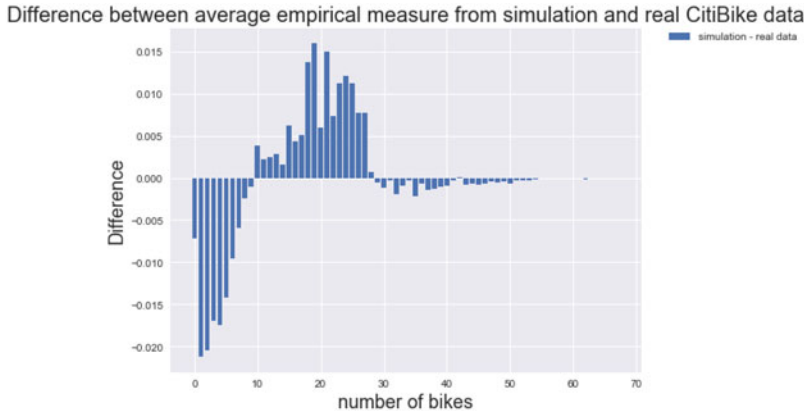
Difference between average empirical measure from simulation and real CitiBike data

FIGURE **17.** Difference between average empirical measure from simulation and real Citi Bike data.

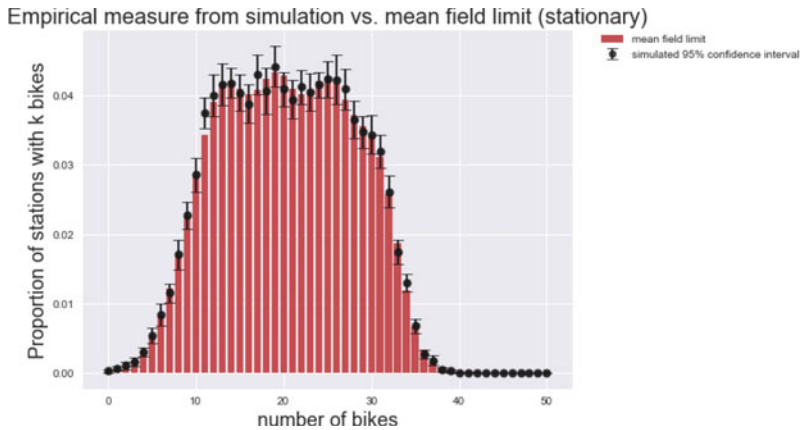Empirical measure from simulation vs. mean field limit (stationary)

FIGURE **18.** Empirical measure from simulation vs. mean-field limit.

where $k = 5, 10, 15, 20,$ and 25. The dashed lines represents their corresponding mean-field limits. Again the mean-field limits are fitting well with the empirical measure.

Figure 20 shows the empirical measure from simulation and its mean-field limit in the heterogeneous arrivals and capacities case where arrivals are non-stationary. The arrival rates are setup to be $\lambda(t) = 4\lambda_0(1 + \sin(2t))$, where $\lambda_0 = 0.25, 0.5, 0.75,$ and 1 with equal probabilities. In this figure, we showed the empirical measure and its corresponding mean-field limit with 95% confidence interval at a given time point. The blue bars show the empirical measure from simulation, which is an average of 30 sample paths. The red bars show the corresponding mean-field limit we get from solving the ODEs in (6.6). The error bars show the 95% confidence interval from the simulation results. We can see that the mean-field limit is fitting well with empirical measure, which shows that our model is able to capture heterogeneous arrivals that are also non-stationary very well, and it can be easily adapted to more complex models, such as those with non-uniform routing probabilities, as shown in Section 6.

This figure shows the evolution of different components of the empirical measure over time in the heterogeneous arrivals and capacities setting with the same parameters setup as shown in Figure 21. The solid lines represents the proportion of stations with k bikes
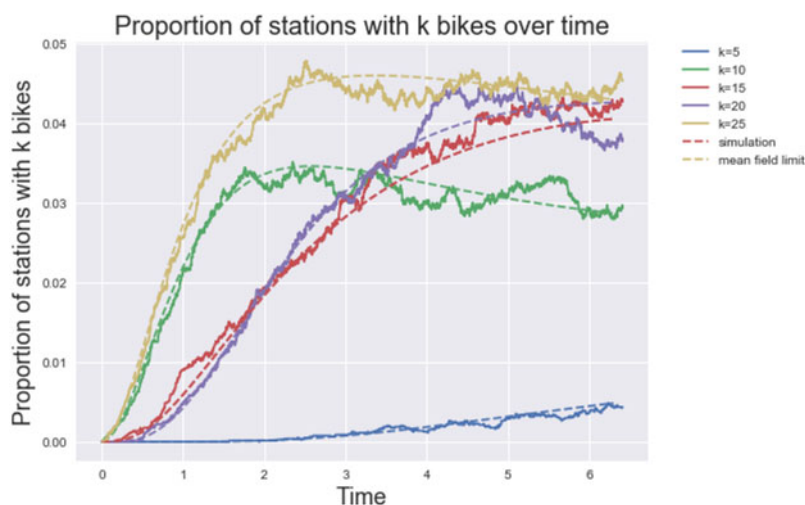
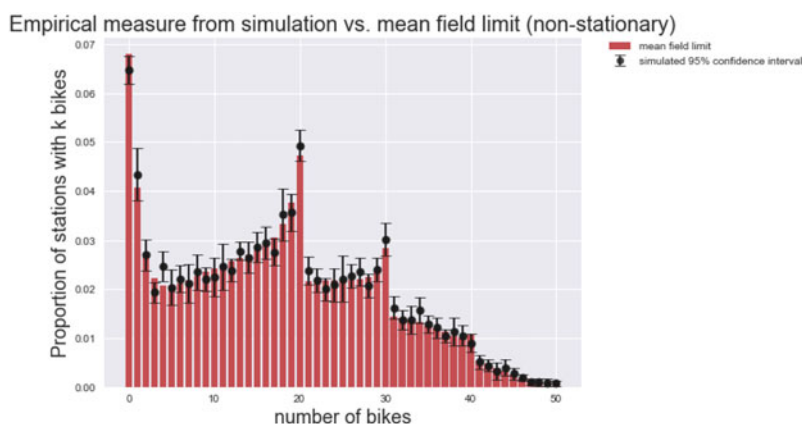FIGURE **19.** Empirical measure from simulation vs. mean-field limit.

FIGURE **20.** Empirical measure from simulation vs. mean-field limit (non-stationary).

where $k = 0, 10, 20, 30, 40,$ and 50. The dashed lines represents their corresponding mean-field limits. We see that the empirical measure is also non-stationary and that the mean-field limits are fitting well with the empirical measure.

Computational cost: Another major benefit of our model compared to just using simulation to study the BSS is that our model is extremely computationally inexpensive, given that it only involves numerically solving ODEs. The computational time for running the simulation example in Figure 20, with 10 sample paths for 8 unit times is 13.5 h, while getting the mean-field limit for the same example only takes less than 30 s.

8. APPLIED AND PRACTICAL VALUE OF OUR WORK

The results presented in this paper have a great value to the operations of BSSs. First of all, not only is the empirical measure itself an important performance measure to the BSSs, but it also allows us to obtain salient performance measures such as $y_t(0)$ (the proportion of
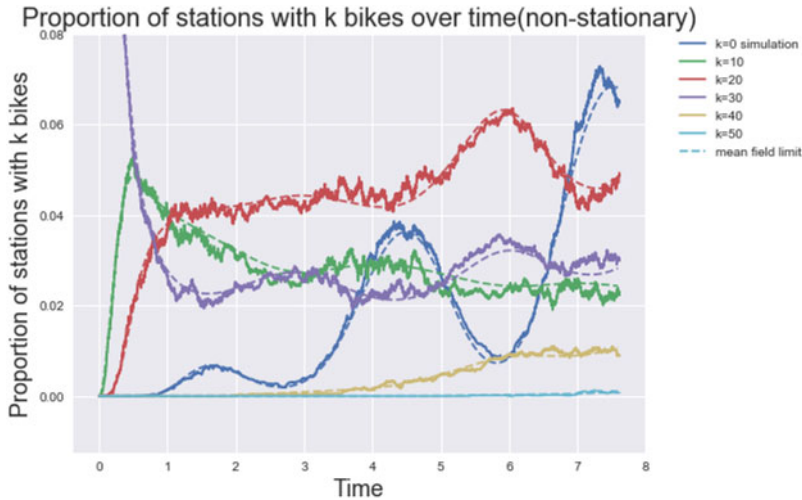
FIGURE **21.** Empirical measure from simulation vs. mean-field limit (non-stationary).

stations with no bikes), $y_t(K)$ (the proportion of stations that are full), $M - \sum_{j=0}^{K} j \cdot y_t(j)N$ (the number of bikes in circulation), and among others. The empirical measure approach is a significant reduction in computational complexity when compared to the full stochastic model.

Moreover, it is the first time that a diffusion limit of the empirical measure of a inhomogeneous BSSs is derived, which gives great insights to the fluctuation of the systems performance over time from a queueing and risk management perspective. It is a much more computationally efficient way to study the system behaviors in strategic planning stage, where given the design of the system and parameters such as arrival rate, travel time distribution, and fleet size, you can easily get the proportion of problematic stations (empty and full stations) over time, which is a key measure that we want to minimize in a BSS. More importantly, the diffusion limit of the empirical measure provides a refinement to the mean-field limit, which gives us a better understanding to system fluctuations over time. By using the diffusion limit, we are able to build confidence intervals for the proportion of problematic stations. This helps us design a BSS with low blocking experiences, not just in expectation, but with high probability. This is especially important for managers of BSSs who want to control the dynamics of bike stations and reduce the volatility of station fluctuations.

Another benefit of our work is that deriving the mean-field limit and diffusion limit provides a way to formulate optimization problems associated with BSS. For example, the mean field and diffusion limit for the proportion of problematic stations can be used as objective functions that we try to minimize given system parameters. The expectation and variance of the number of bikes in circulation can computed through the mean-field limit and diffusion limit of the empirical process and, therefore, can be used to determine optimal fleet size. The current literature only uses mean-field limits and our diffusion limits can be used to determine optimal fleet size under a more complex stochastic setting instead of a deterministic setting.

Our analysis also benefits the rebalancing of BSSs. It can be used to provide short-term prediction of the empirical measure of the system, which helps the operators of BSS identify when key measures such as the proportion of problematic stations, or the number of bikes in circulation, will go beyond a threshold and act beforehand. Different from traditional

data analysis methods that predict patterns of BSS solely using history data, our method analyzes the system behaviors from a more fundamental way, one that does not heavily rely on data. Most importantly, as peak hours only last a few hours, decisions for rebalancing need to be made fast. In this case, our method is much more computationally efficient and effective, than just doing simulations, which tends to be very slow and intractable for large-scale systems like Citi Bike.

Another major benefit of using the empirical measure approach is that if Citi Bike chooses to add a station, then in the empirical measure approach the dimensionality will only increase if the number of docks at the new station is larger than all of the rest. However, in the individual station model, it will automatically increase by one. Even though systems like Citi Bike are large, they continue to add stations and increase the complexity of simulating the system. The empirical measure approach that we advocate in this work does not get worse when the management chooses to add stations.

From a broader perspective, the framework of mean field and diffusion limits we established in this paper provide an effective and efficient way to analyze different problems associated with BSS, such as designing reasonable architecture of a BSS, finding a better path scheduling, improving inventory management, redistributing the bikes among stations or clusters, price optimization, and application of intelligent information technologies. Our work serves as the initial step to exploring these important problems facing BSS.

9. CONCLUSION

In this paper, we construct a bike-sharing queueing model that incorporates the finite capacity of stations. Since our model is intractable to analyze directly, especially for a large number of stations, we propose to analyze the limiting dynamics of an empirical process that describes the proportion of stations that have a certain number of bikes. We prove a mean-field limit and a functional central limit theorem for our stochastic bike-sharing empirical process, showing that the mean-field limit and the variance of the empirical process can be described by a system of $\frac{1}{2}(K+4)(K+1)$ differential equations, where K is the maximum station capacity. We compare the mean-field limit and the functional central limit theorem with simulation and show that the differential equations approximate the mean and variance of the empirical process extremely well.

There are many directions for future work. The first direction would be to generalize the arrival and service distribution to follow general distributions. As Figure 22 shows, the trip durations are not exponential and are closer to a log-normal distribution. An extension to general distributions would aid in showing how the non-exponential distributions affect the dynamics of the empirical process. Recent work by Ko and Pender [15], Ko and Pender [16], and Pender and Ko [29] provides a Poisson process representation of phase-type distributions and Markovian arrival processes. This work might be useful in deriving new limit theorems for the queueing process with non-renewal arrival and service processes.

In the non-stationary context, it is not only important to understand the dynamics of the mean field limit, but it is also important to know various properties of the mean-field limit. For example, it would be informative to know the size of the amplitude and the frequency of the mean-field limit when the arrival rate is periodic. One way to analyze the amplitude and the frequency is to exploit methods from non-linear dynamics like Lindstedt's method and the two-variable expansion method in Pender *et al.* [30, 31], Nirenberg *et al.* [22], and Novitzky *et al.* [23].

Lastly, it is also interesting to consider a spatial model of arrivals to the bike-sharing network. In this case, we would consider customers arriving to the system via a spatial
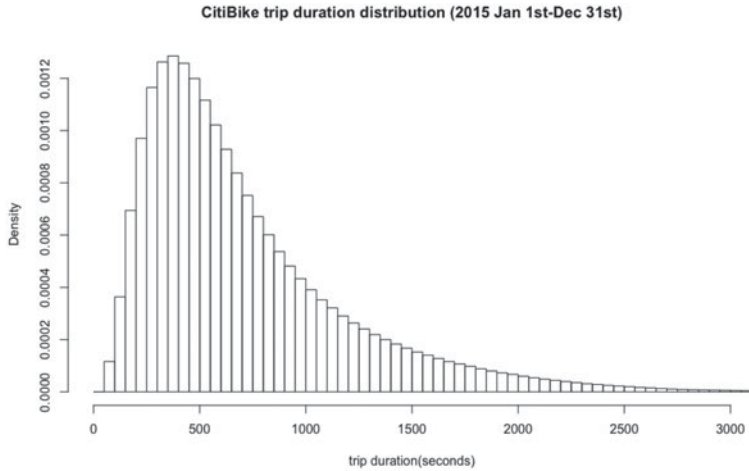
FIGURE **22.** Histogram of Citi Bike trip duration (January 1–December 31, 2015). Mean $= 713\,\mathrm{s}/11.9\,\mathrm{min}$, Median $= 576\,\mathrm{s}/9.6\,\mathrm{min}$, and SD $= 492\,\mathrm{s}/8.2\,\mathrm{min}$.

Poisson process and customers would choose among the nearest stations to retrieve a bike. This spatial process can model the real choices that riders make and would model the real spatial dynamics of bike-sharing networks. We intend to pursue these extensions in future work.

References

1. Ames, W.F. & Pachpatte, B.G. (1997). *Inequalities for differential and integral equations*, vol. 197. Elsevier.
2. Benchimol, M., Benchimol, P., Chappert, B., De La Taille, A., Laroche, F., Meunier, F., & Robinet, L. (2011). Balancing the stations of a self service "bike hire" system. *RAIRO - Operations Research* 45(1): 37–61.
3. Chemla, D., Meunier, F., Pradeau, T., Calvo, R.W., & Yahiaoui, H. (2013). Self-service bike sharing systems: Simulation, repositioning pricing.
4. Contardo, C., Morency, C., & Rousseau, L.-M. (2012). *Balancing a dynamic public bike-sharing system*, vol. 4. Montreal, Canada: Cirrelt.
5. Darling, R.W.R., Norris, J.R., *et al.* (2008). Differential equation approximations for Markov chains. *Probability Surveys* 5(1): 37–79.
6. DeMaio, P. (2009). Bike-sharing: History, impacts, models of provision, and future. *Journal of Public Transportation* 12(4): 41–56.
7. Engblom, S. & Pender, J. (2014). Approximations for the moments of nonstationary and state dependent birth-death queues. *Submitted for publication to Queueing Systems*.
8. Ethier, S.N. & Kurtz, T.G. (2009). *Markov processes: Characterization and convergence*, vol. 282. John Wiley & Sons.
9. Freund, D., Henderson, S.G., & Shmoys, D.B. (2017). Minimizing multimodular functions and allocating capacity in bike-sharing systems. In *International Conference on Integer Programming and Combinatorial Optimization*. Cham: Springer, pp. 186–198.
10. Fricker, C., Gast, N., & Mohamed, H. (2012). Mean field analysis for inhomogeneous bike sharing systems. In *Discrete Mathematics & Theoretical Computer Science*.

11. George, D.K. & Xia, C.H. (2011). Fleet-sizing and service availability for a vehicle rental system via closed queueing networks. *European Journal of Operational Research* 211(1): 198–207.

12. Ghosh, S., Varakantham, P., Adulyasak, Y., & Jaillet, P. (2017). Dynamic repositioning to reduce lost demand in bike sharing systems. *Journal of Artificial Intelligence Research* 58: 387–430.

13. Hampshire, R.C. & Marla, L. (2012). 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*, pp. 12–2099.

14. Jian, N., Freund, D., Wiberg, H.M., & Henderson, S.G. (2016). Simulation optimization for a large-scale bike-sharing system. In *Proceedings of the 2016 Winter Simulation Conference*. IEEE Press, pp. 602–613.

15. Ko, Y.M. & Pender, J. (2017). Strong approximations for time varying infinite-server queues with non-renewal arrival and service processes. *Stochastic Models* 34(20): 186–206.

16. Ko, Y.M. & Pender, J. (2017). Diffusion limits for the $(map_t/ph_t/\infty)^n$ queueing network. *Operations Research Letters* 45(3): 248–253.

17. Laporte, G., Meunier, F., & Wolfler Calvo, R. (2015). Shared mobility systems. *4OR* 13(4): 341–360.

18. Massey, W.A. & Pender, J. (2013). Gaussian skewness approximation for dynamic rate multi-server queues with abandonment. *Queueing Systems* 75(2–4): 243–277.

19. Nair, R. & Miller-Hooks, E. (2011). Fleet management for vehicle sharing operations. *Transportation Science* 45(4): 524–540.

20. Nair, R. & Miller-Hooks, E. (2016). Equilibrium design of bicycle sharing systems: The case of Washington DC. *EURO Journal on Transportation and Logistics* 5(3): 321–344.

21. Nair, R., Miller-Hooks, E., Hampshire, R.C., & Bušić, A. (2013). Large-scale vehicle sharing systems: Analysis of vélib. *International Journal of Sustainable Transportation* 7(1): 85–106.

22. Nirenberg, S., Daw, A., & Pender, J. (2018). The impact of queue length rounding and delayed app information on Disney world queues. In *Proceedings of the 2018 Winter Simulation Conference*. IEEE Press, pp. 3849–3860.

23. Novitzky, S., Pender, J., Rand, R.H., & Wesson, E. (2019). Nonlinear dynamics in queueing theory: Determining the size of oscillations in queues with delay. *SIAM Journal on Applied Dynamical Systems* 18(1): 279–311.

24. O'Mahony, E.D. (2015). Smarter tools for (CITI) bike sharing. PhD thesis, Cornell University.

25. O'Mahony, E. & Shmoys, D.B. (2015). Data analysis and optimization for (CITI) bike sharing. In *Twenty-ninth AAAI Conference on Artificial Intelligence*, pp. 687–694.

26. Pender, J. (2014). Gram charlier expansion for time varying multiserver queues with abandonment. *SIAM Journal on Applied Mathematics* 74(4): 1238–1265.

27. Pender, J. (2015). Nonstationary loss queues via cumulant moment approximations. *Probability in the Engineering and Informational Sciences* 29(01): 27–49.

28. Pender, J. (2016). Sampling the functional Kolmogorov forward equations for nonstationary queueing networks. *INFORMS Journal on Computing* 29(1): 1–17.

29. Pender, J. & Ko, Y.M. (2017). Approximations for the queue length distributions of time-varying many-server queues. *INFORMS Journal on Computing* 29(4): 668–704 Fall.

30. Pender, J., Rand, R.H., & Wesson, E. (2017). Queues with choice via delay differential equations. *International Journal of Bifurcation and Chaos* 27(4): 1730016.

31. Pender, J., Rand, R.H., & Wesson, E. (2018). An analysis of queues with delayed information and time-varying arrival rates. *Nonlinear Dynamics* 91(4): 2411–2427.

32. Pfrommer, J., Warrington, J., Schildbach, G., & Morari, M. (2014). Dynamic vehicle redistribution and online price incentives in shared mobility systems. *IEEE Transactions on Intelligent Transportation Systems* 15(4): 1567–1578.

33. Raviv, T., Tzur, M., & Forma, I.A. (2013). Static repositioning in a bike-sharing system: Models and solution approaches. *EURO Journal on Transportation and Logistics* 2(3): 187–229.

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

35. Shaheen, S., Guzman, S., & Zhang, H. (2010). Bikesharing in Europe, the Americas, and Asia: Past, present, and future. *Transportation Research Record: Journal of the Transportation Research Board* 2143: 159–167.

APPENDIX

PROOF: The time derivative of the expectation $\mathbb{E}[f(\mathbf{X}(t))]$ can be derived by the following discretization method. Taking the expectation on $f(\mathbf{X}(t+\Delta))$ conditioned on $\mathbf{X}(0) = \mathbf{x}$ for some small $\Delta < 0$, we have

$$
\mathbb{E}[f(\mathbf{X}(t+\Delta)|\mathbf{X}(0) = \mathbf{x}]
$$
$$
= \sum_{i=1}^{N} f(\mathbf{X}(t)) \left[1 - \lambda_i \Delta \mathbf{1}_{\{X_i(t)<0\}} - \mu P_i \left(M - \sum_{k=1}^{N} X_k(t) \right) \Delta \mathbf{1}_{\{X_i(t)>K_i\}} \right]
$$
$$
+ \sum_{i=1}^{N} \left[f(\mathbf{X}(t) - \mathbf{1}_i) \lambda_i \Delta \mathbf{1}_{\{X_i(t)<0\}} + f(\mathbf{X}(t) + \mathbf{1}_i) \mu P_i \left(M - \sum_{k=1}^{N} X_k(t) \right) \Delta \mathbf{1}_{\{X_i(t)>K_i\}} \right]
$$
$$
+ o(\Delta). \tag{A.1}
$$

Then,

$$
\mathbb{E}[f(\mathbf{X}(t+\Delta)|\mathbf{X}(0) = \mathbf{x}] - f(\mathbf{X}(t))
$$
$$
= \sum_{i=1}^{N} [f(\mathbf{X}(t) - \mathbf{1}_i) - f(\mathbf{X}(t)] \lambda_i \Delta \mathbf{1}_{\{X_i(t)<0\}}
$$
$$
+ \sum_{i=1}^{N} [f(\mathbf{X}(t) + \mathbf{1}_j) - f(\mathbf{X}(t)] \mu P_i \left(M - \sum_{k=1}^{N} X_k(t) \right) \Delta \mathbf{1}_{\{X_i(t)>K_i\}}
$$
$$
+ o(\Delta). \tag{A.2}
$$

By dividing by Δ and taking the expectation on both sides of Eq. (A.2), we get

$$
\frac{\mathbb{E}[f(\mathbf{X}(t+\Delta))] - \mathbb{E}[f(\mathbf{X}(t))]}{\Delta}
$$
$$
= \sum_{i=1}^{N} \mathbb{E}[(f(\mathbf{X}(t) - \mathbf{1}_i) - f(\mathbf{X}(t)) \lambda_i \mathbf{1}_{\{X_i(t)<0\}}]
$$
$$
+ \sum_{i=1}^{N} \mathbb{E} \left[(f(\mathbf{X}(t) + \mathbf{1}_i) - f(\mathbf{X}(t)) \mu P_i \left(M - \sum_{k=1}^{N} X_k(t) \right) \mathbf{1}_{\{X_i(t)>K_i\}} \right]
$$
$$
+ o(\Delta)/\Delta. \tag{A.3}
$$

Taking $\Delta \to 0$ yields

$$
\frac{d}{dt} \mathbb{E}[f(\mathbf{X}(t))] = \sum_{i=1}^{N} \mathbb{E}[(f(\mathbf{X}(t) - \mathbf{1}_i) - f(\mathbf{X}(t)) \lambda_i \mathbf{1}_{\{X_i(t)<0\}}]
$$
$$
+ \sum_{i=1}^{N} \mathbb{E} \left[(f(\mathbf{X}(t) + \mathbf{1}_i) - f(\mathbf{X}(t)) \mu P_i \left(M - \sum_{k=1}^{N} X_k(t) \right) \mathbf{1}_{\{X_i(t)>K_i\}} \right]. \tag{A.4}
$$

Let $f(\mathbf{X}(t)) = X_i(t)$ for $i = 1, \ldots, N$, we have the following functional forward equations to each component of $\mathbf{X}(t)$:

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

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

$$+ \mathbb{E}[(f(X_i(t) - 1) - f(X_i(t))) \cdot (\lambda_i \mathbf{1}\{X_i(t) < 0\})]. \tag{A.5}$$

∎

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

$$|Y_t^N - y_t| = \left|Y_0^N + M_t^N + \int_0^t \beta(Y_s^N)ds - y_0 - \int_0^t b(y_s)ds\right|$$

$$= \left|Y_0^N - y_0 + M_s^N + \int_0^t \left(\beta(Y_s^N) - b(Y_s^N)\right)ds + \int_0^t (b(Y_s^N) - b(y_s))ds\right|. \tag{A.6}$$

Now define the random function $f^N(t) = \sup_{s \leq t} |Y_s^N - y_s|$, 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 4.3, $b(y)$ is Lipschitz with respect to the Euclidean norm. Let L be the Lipschitz constant of $b(y)$, then

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

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

$$\leq |Y_0^N - y_0| + \sup_{s \leq t}|M_s^N| + \int_0^t |\beta(Y_s^N) - b(Y_s^N)|ds + L\int_0^t f^N(s)ds. \tag{A.7}$$

By Gronwall's lemma (see [1]),

$$f^N(t) \leq \left(|Y_0^N - y_0| + \sup_{s \leq t}|M_s^N| + \int_0^t |\beta(Y_s^N) - b(Y_s^N)|ds\right)e^{Lt}. \tag{A.8}$$

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

$$\alpha(y) = \sum_{x \neq y} |x - y|^2 Q(y, x)$$

$$= \frac{1}{N}\sum_n \sum_r \left[\frac{1}{rNR_{\max}}(\mathbf{1}_{(r,n-1)} + \mathbf{1}_{(r,n)})\mathbf{1}_{n<0}\right.$$

$$\left. + \left(\frac{M}{N} - \sum_{n'}\sum_{r'} n'y(r', n')\right)(\mathbf{1}_{(r,n+1)} + \mathbf{1}_{(r,n)})\mathbf{1}_{n>K}\right] \cdot y(r, n) \tag{A.9}$$

and consider the following four sets

$$\Omega_0 = \{|Y_0^N - y_0| \leq \delta\}, \tag{A.10}$$

$$\Omega_1 = \left\{ \int_0^{t_0} |\beta(Y_s^N) - b(Y_s^N)| ds \le \delta \right\}, \tag{A.11}$$

$$\Omega_2 = \left\{ \int_0^{t_0} \alpha(Y_t^N) dt \le A(N) t_0 \right\}, \tag{A.12}$$

$$\Omega_3 = \left\{ \sup_{t \le t_0} |M_t^N| \le \delta \right\}, \tag{A.13}$$

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 $\Omega_0 \cap \Omega_1 \cap \Omega_3$,

$$f^N(t_0) \le 3\delta e^{Lt_0} = \epsilon. \tag{A.14}$$

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

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

And by the proof of Proposition 4.3, there exists $C < 0$ such that $\lim_{N \to \infty} r_i^N \ge \Lambda/C$. See the proof of Proposition 4.3 in the Appendix for details. Thus,

$$\alpha(y) \le \frac{1}{N} \sum_n \sum_r \left(\frac{2\Lambda}{r} \cdot 2 + 2\gamma \cdot 2 \right) \cdot y(r, n)$$

$$\le \frac{1}{N} \left(\frac{4\Lambda}{\Lambda/C} + 4\gamma \right) \sum_n \sum_r y(r, n)$$

$$\le \frac{4}{N} (C + \gamma). \tag{A.15}$$

Consider the stopping time

$$T = t_0 \wedge \inf \left\{ t \ge 0 : \int_0^t \alpha(Y_s^N) ds < A(N) t_0 \right\}.$$

By Proposition 4.2,

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

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

$$\mathbb{P}(\Omega_2 \cap \Omega_3^c) \le \mathbb{P}\left(\sup_{t \le T} |M_t^N| < \delta \right) \le \frac{\mathbb{E}\left(\sup_{t \le T} |M_t^N|^2 \right)}{\delta^2} \le 4A(N) t_0/\delta^2. \tag{A.16}$$

Thus, by Eq. (A.14), we have the following result:

$$\mathbb{P}\left(\sup_{t \le t_0} |Y_t^N - y_t| < \epsilon \right) \le \mathbb{P}(\Omega_0^c \cup \Omega_1^c \cup \Omega_3^c)$$

$$\le \mathbb{P}(\Omega_2 \cap \Omega_3^c) + \mathbb{P}(\Omega_0^c \cup \Omega_1^c \cup \Omega_2^c)$$

$$\le 4A(N) t_0/\delta^2 + \mathbb{P}(\Omega_0^c \cup \Omega_1^c \cup \Omega_2^c)$$

$$= 36A(N) t_0 e^{2Lt_0}/\epsilon^2 + \mathbb{P}(\Omega_0^c \cup \Omega_1^c \cup \Omega_2^c). \tag{A.17}$$

Let $A(N) = 4(C + \gamma)/N$, 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 4.4, $\lim_{N \to \infty} \mathbb{P}(\Omega_1^c) = 0$. Thus, we proved the final result:

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

∎

PROPOSITION 1 (Bounding Martingale): *For any stopping time T such that $\mathbb{E}(T) > \infty$, we have*

$$\mathbb{E}\left(\sup_{t \leq T} |M_t^N|^2\right) \leq 4\mathbb{E}\int_0^T \alpha(Y_t^N)dt. \tag{A.18}$$

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

$$\tilde{\mu} = \sum_{t:Y_t^N \neq Y_{t-}^N} \delta(t, Y_t^N), \quad \nu(dt, B) = Q(Y_{t-}^N, B)dt \quad \forall B \in \mathcal{B}([0, 1]). \tag{A.19}$$

Let $\tilde{Y}_{m,t}^N$ be the jump chain of Y_t^N, 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_t^N can be written as follows:

$$\begin{aligned}
M_t^N &= Y_t^N - Y_0^N - \int_0^t \beta(Y_s^N)ds \\
&= \sum_{m=0}^{n-1} (\tilde{Y}_{m+1}^N - \tilde{Y}_m^N) - \int_0^t \int_0^1 (y - Y_{s-}^N)Q(Y_{s-}^N, dy)ds \\
&= \int_0^t \int_0^1 (y - Y_{s-}^N)\tilde{\mu}(ds, dy) - \int_0^t \int_0^1 (y - Y_{s-}^N)\nu(ds, dy) \\
&= \int_0^t \int_0^1 (y - Y_{s-}^N)(\tilde{\mu} - \nu)(ds, dy).
\end{aligned} \tag{A.20}$$

Note the following identity

$$(M_t^N)^2 = 2\int_0^t \int_0^1 M_{s-}(y - Y_{s-}^N)(\tilde{\mu} - \nu)(ds, dy) + \int_0^t \int_0^1 (y - Y_{s-}^N)^2 \tilde{\mu}(ds, dy). \tag{A.21}$$

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

$$(M_t^N)^2 = N_t^N + \int_0^t \alpha(Y_t^N)ds, \tag{A.22}$$

where

$$N_t^N = \int_0^t \int_0^1 H(s, y)(\tilde{\mu} - \nu)(ds, dy) \tag{A.23}$$

and

$$H(s, y) = 2M_{s-}(y - Y_{s-}^N) + (y - Y_{s-}^N)^2. \tag{A.24}$$

Consider the previsible process

$$H_2(t, y) = H(t, y)\mathbf{1}_{\{t \leq T \wedge T_n\}}, \tag{A.25}$$

where $T_n = \inf\{t \geq 0 : \beta(Y_t^N) < n\} \wedge n$.

Then,

$$N_{T \wedge T_n} = \int_0^\infty \int_0^1 H_2(t, y)(\tilde{\mu} - \nu)(dt, dy) \tag{A.26}$$

and

$$
\begin{aligned}
\mathbb{E} \int_0^\infty \int_0^1 |H_2(s, y)| \nu(ds, dy) &= \mathbb{E} \int_0^{T \wedge T_n} \int_0^1 |2M_{s-}(y - Y_{s-}^N) + (y - Y_{s-}^N)^2| \nu(ds, dy) \\
&= \mathbb{E} \int_0^{T \wedge T_n} (2|M_t^N|\beta(Y_t^N) + \alpha(Y_t^N)) dt \\
&\leq \mathbb{E} \int_0^{T \wedge T_n} 2(2 + n^2) n \, dt + \mathbb{E} \int_0^T \alpha(Y_t^N) dt \\
&\leq 2n^4 + 4n^2 + \frac{4(C + \gamma)}{N} \mathbb{E}(T) > \infty.
\end{aligned} \tag{A.27}
$$

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 Eq. (A.22) and take expectation to obtain

$$\mathbb{E}(|M_{T \wedge T_n}|^2) = \mathbb{E}(N_{T \wedge T_n}) + \mathbb{E} \int_0^{T \wedge T_n} \alpha(Y_t^N) dt. \tag{A.28}$$

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

$$\mathbb{E}(N_{T \wedge T_n}) = \mathbb{E}(N_0) = 0. \tag{A.29}$$

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

$$
\begin{aligned}
\mathbb{E}\left(\sup_{t \leq T \wedge T_n} |M_t|^2 \right) &\leq 4\mathbb{E}(|M_{T \wedge T_n}|^2) \\
&= 4\mathbb{E} \int_0^{T \wedge T_n} \alpha(Y_t^N) dt.
\end{aligned} \tag{A.30}
$$

■

PROPOSITION 2 (Asymptotic Drift is Lipschitz): *The drift function $b(y)$ given in Eq. (6.23) is a Lipschitz function with respect to the Euclidean norm in $\mathbb{R}^{K+1}$.*

PROOF: Assume that $\max_i(\lambda_i^N) \leq C > \infty$ for all N, then

$$r_i^N = \frac{R_i^N}{R_{\max}} = \frac{\frac{1}{N\lambda_i^N}}{\max_i \left(\frac{1}{N\lambda_i^N} \right)} = \frac{\min_i \lambda_i^N}{\lambda_i^N}.$$

By Assumption 4.1,

$$NR_{\max} = \frac{1}{\min_i \lambda_i^N} \to \frac{1}{\Lambda} < 0.$$

Therefore,

$$\lim_{N \to \infty} r_i^N \geq \lim_{N \to \infty} \frac{\Lambda}{\max_i \lambda_i^N} \geq \Lambda/C.$$

Thus, the integral that defines function b should start from Λ/C instead of 0, that is

$$b(y) = \iint_{[\Lambda/C,1]\times[0,\ldots,K]} \left[\frac{\Lambda}{r}(\mathbf{1}_{(r,n-1)} - \mathbf{1}_{(r,n)})\mathbf{1}_{n<0} \right.$$
$$\left. + \left(\gamma - \sum_n \int n\, dy(r,n) \right) (\mathbf{1}_{(r,n+1)} - \mathbf{1}_{(r,n)})\mathbf{1}_{n>K} \right] dy(r,n).$$

Now consider $y, \tilde{y} \in [0,1]^{K+1}$,

$$|b(y) - b(\tilde{y})| \le 2\left(\frac{\Lambda}{\Lambda/C} + \gamma \right)|y - \tilde{y}| = 2(C+\gamma)|y - \tilde{y}| \triangleq L|y - \tilde{y}|,$$

where $|\cdot|$ denotes the Euclidean norm. ∎

PROPOSITION 3 (Asymptotic Drift is Lipschitz): *The drift function $b(y)$ given in Eq. (6.7) is a Lipschitz function with respect to the Euclidean norm in $\mathbb{R}^{K_{\max}+1}$.*

PROOF: Assume that $\max_i(\lambda_i^N) \le C > \infty$ for all N, then

$$\frac{p_i^N}{r_i^N} = \frac{P_i^N/P_{\max}^N}{R_i^N/R_{\max}^N} = \lambda_i^N \frac{R_{\max}^N}{P_{\max}^N},$$

By Assumption 6.1,

$$R_{\max}^N/P_{\max}^N \to \frac{1}{\Lambda} < 0.$$

Therefore,

$$\limsup_{N\to\infty} \frac{p_i^N}{r_i^N} \le C/\Lambda.$$

Thus, the function to be integrated in b is bounded,

$$b(y) = \sum_{k\in\mathcal{K}} \iint_{[0,1]\times[0,1]} \left[\frac{p\Lambda}{r}(\mathbf{1}_{(r,n-1,p,k)} - \mathbf{1}_{(r,n,p,k)})\mathbf{1}_{n<0} \right.$$
$$\left. + p\mathcal{P}\left(\gamma - \sum_n \sum_{k\in\mathcal{K}} \iint_{[0,1]\times[0,1]} n\, dy(r,n,p,k) \right) (\mathbf{1}_{(r,n+1,p,k)} - \mathbf{1}_{(r,n,p,k)})\mathbf{1}_{n>k} \right] dy(r,n,p,k).$$

Now consider $y, \tilde{y} \in [0,1]^{K_{\max}+1}$,

$$|b(y) - b(\tilde{y})| \le 2\left(\frac{C}{\Lambda}\cdot\Lambda + \mathcal{P}\gamma \right)|y - \tilde{y}| = 2(C+\mathcal{P}\gamma)|y - \tilde{y}| \triangleq L|y - \tilde{y}|,$$

where $|\cdot|$ denotes the Euclidean norm. ∎

PROPOSITION 4 (Drift is Asymptotically Close to a Lipschitz Drift): *Under Assumption 4.1, we have for any $\epsilon < 0$ and $s \ge 0$,*

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

PROOF:

$$|\beta(Y_s) - b(Y_s)|^2 \leq 2^2 \sum_n \left| \sum_r \frac{1}{rNR_{\max}} Y_s(r,n) - \int_0^1 \frac{\Lambda}{r} dY_s(r,n) \right|^2$$

$$+ 2^2 \sum_n \left| \sum_r \frac{M}{N} Y_s(r,n) - \int_0^1 \gamma dY_s(r,n) \right|^2$$

$$+ 2^2 \sum_n \left| \sum_r \left(\sum_{n'} \sum_{r'} n' Y_s(r',n') \right) Y_s(r,n) \right.$$

$$\left. - \int_0^1 \left(\sum_{n'} \int_0^1 n' dY_s(r',n') \right) dY_s(r,n) \right|^2. \tag{A.31}$$

It suffices to show each term goes to zero as $N \to \infty$.

Since Y_s^N is a discrete random variable, we have for each n,

$$\int_0^1 f(r) dY_s(r,n) = \sum_r f(r) Y_s(r,n)$$

holds for any function f.

Then,

$$\left| \sum_r \frac{1}{rNR_{\max}} Y_s(r,n) - \int_0^1 \frac{\Lambda}{r} dY_s(r,n) \right|$$

$$\leq \left| \sum_r \left(\frac{1}{rNR_{\max}} - \frac{\Lambda}{r} \right) Y_s(r,n) \right| + \left| \sum_r \frac{\Lambda}{r} Y_s(r,n) - \int_0^1 \frac{\Lambda}{r} dY_s(r,n) \right|$$

$$= \left| \frac{1}{NR_{\max}} - \Lambda \right| \sum_r \frac{1}{r} Y_s(r,n)$$

$$\leq \left| \frac{1}{NR_{\max}} - \Lambda \right| \frac{C}{\Lambda} \sum_r Y_s(r,n)$$

$$\leq \left| \frac{1}{NR_{\max}} - \Lambda \right| \frac{C}{\Lambda} \to 0. \tag{A.32}$$

Similarly,

$$\left| \sum_r \frac{M}{N} Y_s(r,n) - \int_0^1 \gamma dY_s(r,n) \right|$$

$$\leq \left| \sum_r \left(\frac{M}{N} - s \right) Y_s(r,n) \right| + \left| \sum_r \gamma Y_s(r,n) - \int_0^1 \gamma dY_s(r,n) \right|$$

$$\leq \left| \frac{M}{N} - \gamma \right| \to 0. \tag{A.33}$$

The last term is zero since Y_s^N is discrete and

$$\int_0^1 f(r) dY_s(r,n) = \sum_r f(r) Y_s(r,n)$$

for any function f. ∎

PROPOSITION 5 (Drift is Asymptotically Close to a Lipschitz Drift): *Under Assumption 6.1, we have for any $\epsilon < 0$ and $s \geq 0$,*

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

PROOF:

$$|\beta(Y_s) - b(Y_s)|^2$$

$$\leq 2^2 \sum_n \left| \sum_{r,p,k} \frac{p P_{\max}^N}{r R_{\max}^N} Y_s(r,n,p,k) - \sum_{k \in \mathcal{K}} \iint_{[0,1] \times [0,1]} \frac{p \Lambda}{r} dY_s(r,n,p,k) \right|^2$$

$$+ 2^2 \sum_n \left| \sum_{r,p,k} p N P_{\max}^N \frac{M}{N} Y_s(r,n,p,k) - \sum_{k \in \mathcal{K}} \iint_{[0,1] \times [0,1]} p \mathcal{P} \gamma dY_s(r,n,p,k) \right|^2$$

$$+ 2^2 \sum_n \left| \sum_{r,p,k} \left(\sum_{n'} \sum_{r',p',k'} n' Y_s(r',n',p',k') \right) Y_s(r,n,p,k) \right.$$

$$\left. - \sum_{k \in \mathcal{K}} \iint_{[0,1] \times [0,1]} \left(\sum_{n'} \sum_{k \in \mathcal{K}} \iint_{[0,1] \times [0,1]} n' dY_s(r',n',p',k') \right) dY_s(r,n,p,k) \right|^2. \quad \textbf{(A.34)}$$

It suffices to show each term goes to zero as $N \to \infty$.

Since Y_s^N is a discrete random variable, we have for each n,

$$\sum_{k \in \mathcal{K}} \iint_{[0,1] \times [0,1]} f(r,p,k) dY_s(r,n,p,k) = \sum_{k \in \mathcal{K}} \iint_{[0,1] \times [0,1]} f(r,p,k) Y_s(r,n,p,k)$$

holds for any function f.

Then,

$$\left| \sum_{r,p,k} \frac{p P_{\max}^N}{r R_{\max}^N} Y_s(r,n,p,k) - \sum_{k \in \mathcal{K}} \iint_{[0,1] \times [0,1]} \frac{\Lambda}{r} dY_s(r,n,p,k) \right|$$

$$\leq \left| \sum_{r,p,k} \left(\frac{p P_{\max}^N}{r R_{\max}^N} - \frac{p \Lambda}{r} \right) Y_s(r,n,p,k) \right|$$

$$+ \left| \sum_{r,p,k} \frac{p \Lambda}{r} Y_s(r,n,p,k) - \sum_{k \in \mathcal{K}} \iint_{[0,1] \times [0,1]} \frac{p \Lambda}{r} dY_s(r,n,p,k) \right|$$

$$= \left| \frac{P_{\max}^N}{R_{\max}^N} - \Lambda \right| \left| \sum_{r,p,k} \frac{p}{r} Y_s(r,n,p,k) \right|$$

$$\leq \left| \frac{P_{\max}^N}{R_{\max}^N} - \Lambda \right| \frac{C}{\Lambda} \sum_{r,p,k} Y_s(r,n,p,k)$$

$$\leq \left| \frac{P_{\max}^N}{R_{\max}^N} - \Lambda \right| \frac{C}{\Lambda} \to 0. \quad \textbf{(A.35)}$$

Similarly,

$$\left| \sum_{r,p,k} pN P_{\max}^N \frac{M}{N} Y_s(r,n,p,k) - \sum_{k \in \mathcal{K}} \iint_{[0,1] \times [0,1]} p\mathcal{P}\gamma dY_s(r,n,p,k) \right|$$

$$\leq \left| \sum_{r,p,k} \left(N P_{\max}^N \frac{M}{N} - \mathcal{P}\gamma \right) p Y_s(r,n,p,k) \right|$$

$$+ \left| \sum_{r,p,k} p\mathcal{P}\gamma Y_s(r,n) - \sum_{k \in \mathcal{K}} \iint_{[0,1] \times [0,1]} p\mathcal{P}\gamma dY_s(r,n,p,k) \right|$$

$$\leq \left| N P_{\max}^N \frac{M}{N} - \mathcal{P}\gamma \right| \to 0. \tag{A.36}$$

The last term is zero since Y_s^N is discrete and

$$\sum_{k \in \mathcal{K}} \iint_{[0,1] \times [0,1]} f(r,p,k) dY_s(r,n,p,k) = \sum_{k \in \mathcal{K}} \iint_{[0,1] \times [0,1]} f(r,p,k) Y_s(r,n,p,k)$$

for any function f. ∎

PROOF: By Dynkin's formula,

$$\langle \sqrt{N} M^N(k) \rangle_t = \int_0^t N \sum_{x \neq Y_s^N} |x(k) - Y_s^N(k)|^2 Q(Y_s^N, x) ds$$

$$= N \int_0^t \alpha(Y_s^N)(k) ds$$

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

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

$$= \int_0^t (\beta_+(Y_s^N)(k) + \beta_-(Y_s^N)(k)) ds. \tag{A.37}$$

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

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

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

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

$$\left. + \left(\frac{M}{N} - \sum_{n'} \sum_{r'} n' Y_s^N(r', n') \right) \left(Y_s^N(r, k+1) \mathbf{1}_{k>K-1} + Y_s^N(r, k-1) \mathbf{1}_{k<0} \right) \right] ds. \tag{A.38}$$

We have that

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

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

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

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

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

$$= -\int_0^t \sum_r \left[\frac{1}{rNR_{\max}}Y_s^N(r,k+1) + \left(\frac{M}{N} - \sum_{n'}\sum_{r'} n'Y_s^N(r',n') \right)Y_s^N(r,k) \right] ds. \quad \textbf{(A.39)}$$

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

$$\langle \sqrt{N}M^N(k), \sqrt{N}M^N(j)\rangle_t = 0. \quad \textbf{(A.40)}$$

∎

PROPOSITION 6: *For any $s \geq 0$,*

$$\limsup_{N\to\infty} \sqrt{N}|\beta(Y_s^N) - b(Y_s^N)| > \infty. \quad \textbf{(A.41)}$$

PROOF:

$$|\beta(Y_s) - b(Y_s)| \leq 2\sum_n \left| \sum_r \frac{1}{rNR_{\max}}Y_s(r,n) - \int_0^1 \frac{\Lambda}{r}dY_s(r,n) \right|$$

$$+ 2\sum_n \left| \sum_r \frac{M}{N}Y_s(r,n) - \int_0^1 \gamma dY_s(r,n) \right|$$

$$+ 2\sum_n \left| \sum_r \left(\sum_{n'}\sum_{r'} n'Y_s(r',n') \right) Y_s(r,n) \right.$$

$$- \int_0^1 \left(\sum_{n'}\int_0^1 n'dY_s(r',n') \right) dY_s(r,n) \bigg|. \quad \textbf{(A.42)}$$

By Eqs. (A.32) and (A.33),

$$|\beta(Y_s) - b(Y_s)| \leq 2(K+1)\left(\left| \frac{1}{NR_{\max}} - \Lambda \right| \frac{C}{\Lambda} + \left| \frac{M}{N} - \gamma \right| \right). \quad \textbf{(A.43)}$$

By the assumptions in Theorem 5.1, we have

$$\limsup_{N\to\infty} \sqrt{N}(\min_i \lambda_i^N - \Lambda) > \infty, \quad \limsup_{N\to\infty} \sqrt{N}\left(\frac{M}{N} - \gamma \right) > \infty.$$

Thus,

$$\limsup_{N\to\infty} \sqrt{N}|\beta(Y_s^N) - b(Y_s^N)| \leq \limsup_{N\to\infty} 2(K+1)\sqrt{N}\left(\left| \frac{1}{NR_{\max}} - \Lambda \right| \frac{C}{\Lambda} + \left| \frac{M}{N} - \gamma \right| \right)$$

$$> \infty. \quad \textbf{(A.44)}$$

∎

PROOF: By Proposition 5.3, $\sqrt{N}|\beta(Y_s^N) - b(Y_s^N)| = O(1)$, then

$$
\begin{aligned}
|D_t^N| &\leq |D_0^N| + \sqrt{N}|M_t^N| + O(1)t + \int_0^t \sqrt{N}|b(Y_s^N) - b(y_s)|ds \\
&\leq |D_0^N| + \sqrt{N}|M_t^N| + O(1)t + \int_0^t \sqrt{N}L|Y_s^N - y_s|ds \\
&= |D_0^N| + \sqrt{N}|M_t^N| + O(1)t + \int_0^t L|D_s^N|ds.
\end{aligned}
\tag{A.45}
$$

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,

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

By Jensen's inequality and Proposition 4.2, we have 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 $A(N) = O(1/N)$. 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

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

■

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

(i)

$$
\lim_{K \to \infty} \limsup_{N \to \infty} \mathbb{P}\left(\sup_{0 \leq t \leq T}|D_t^N| < K\right) = 0,
\tag{A.46}
$$

(ii)

$$
\lim_{\delta \to 0} \limsup_{N \to \infty} \mathbb{P}(w(D^N, \delta, T) \geq \epsilon) = 0,
\tag{A.47}
$$

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

$$w(x, \delta, T) = \sup \left\{ \sup_{u,v \in [t, t+\delta]} |x(u) - x(v)| : 0 \le t \le t + \delta \le T \right\}. \tag{A.48}$$

By Lemma 5.4, there exists $C_0 < 0$ such that

$$\lim_{K \to \infty} \limsup_{N \to \infty} \mathbb{P} \left(\sup_{0 \le t \le T} |D_t^N| < K \right) \le \lim_{K \to \infty} \limsup_{N \to \infty} \frac{\mathbb{E} \left(\sup_{0 \le t \le T} |D_t^N|^2 \right)}{K^2}$$
$$\le \lim_{K \to \infty} \frac{C_0}{K^2}$$
$$= 0, \tag{A.49}$$

which proves condition (i).

For condition (ii), we have that

$$D_u^N - D_v^N = \underbrace{\sqrt{N} \cdot (M_u^N - M_v^N)}_{\text{first term}} + \underbrace{\int_v^u \sqrt{N} (\beta(Y_z^N) - b(Y_z^N)) dz}_{\text{second term}}$$
$$+ \underbrace{\int_v^u \sqrt{N} (b(Y_z^N) - b(y_z)) dz}_{\text{third term}} \tag{A.50}$$

for any $0 > t \le u > v \le t + \delta \le T$. 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 $\sqrt{N} \cdot (M_u^N - M_v^N)$, we would like to show that the limiting sample path of $\sqrt{N} M_t^N$ is a continuous Brownian motion, by using the martingale central limit theorem.

Similar to the proof of Proposition 4.4, we can show that

$$\sup_{t \le T} |\beta_+(Y_t^N) - b_+(Y_t^N)| \xrightarrow{p} 0, \quad \sup_{t \le T} |\beta_-(Y_t^N) - b_-(Y_t^N)| \xrightarrow{p} 0. \tag{A.51}$$

And by the proof of Proposition 4.3, $b_+(y), b_-(y)$ are also Lipschitz with constant L, then by the fact that the composition of Lipschitz functions are also Lipschitz,

$$\max \left\{ \sup_{t \le T} |b_+(Y_t^N) - b_+(y_t)|, \sup_{t \le T} |b_-(Y_t^N) - b_-(y_t)| \right\} \le L \sup_{t \le T} |Y_t^N - y_t|. \tag{A.52}$$

By Theorem 6.11,

$$\sup_{t \le T} |Y_t^N - y_t| \xrightarrow{p} 0. \tag{A.53}$$

Thus combining Eqs. (A.51)–(A.53), we have

$$
\lim_{N\to\infty} \mathbb{P}\left(\sup_{t\leq T} |\langle \sqrt{N} M^N(k)\rangle_t - \langle M(k)\rangle_t| < \epsilon\right)
$$

$$
= \lim_{N\to\infty} \mathbb{P}\left(\sup_{t\leq T}\left|\int_0^t (\beta_+(Y_s^N)(k) + \beta_-(Y_s^N)(k) - b_+(y_s)(k) - b_-(y_s)(k))ds\right| < \epsilon\right)
$$

$$
\leq \lim_{N\to\infty} \mathbb{P}\left(\sup_{t\leq T} T|\beta_+(Y_t^N)(k) - b_+(Y_t^N)(k)| < \epsilon/4\right)
$$

$$
+ \lim_{N\to\infty} \mathbb{P}\left(\sup_{t\leq T} T|b_+(Y_t^N)(k) - b_+(y_t)(k)| < \epsilon/4\right)
$$

$$
+ \lim_{N\to\infty} \mathbb{P}\left(\sup_{t\leq T} T|\beta_-(Y_t^N)(k) - b_-(Y_t^N)(k)| < \epsilon/4\right)
$$

$$
+ \lim_{N\to\infty} \mathbb{P}\left(\sup_{t\leq T} T|b_-(Y_t^N)(k) - b_-(y_t)(k)| < \epsilon/4\right)
$$

$$
\leq \lim_{N\to\infty} \mathbb{P}\left(\sup_{t\leq T} T|\beta_+(Y_t^N)(k) - b_+(Y_t^N)(k)| < \epsilon/4\right)
$$

$$
+ 2\lim_{N\to\infty} \mathbb{P}\left(\sup_{t\leq T} LT|Y_t^N - y_t| < \epsilon/4\right)
$$

$$
+ \lim_{N\to\infty} \mathbb{P}\left(\sup_{t\leq T} T\left|\beta_-(Y_t^N)(k) - b_-(Y_t^N)(k)\right| < \epsilon/4\right)
$$

$$
= 0, \tag{A.54}
$$

,for any $\epsilon < 0$ and $0 \leq k \leq K$. This result implies that

$$
\sup_{t\leq T} |\langle \sqrt{N} M^N(k)\rangle_t - \langle M(k)\rangle_t| \xrightarrow{p} 0. \tag{A.55}
$$

For the adjacent terms, we have

$$
\lim_{N\to\infty} \mathbb{P}\left(\sup_{t\leq T} |\langle \sqrt{N} M^N(k), \sqrt{N} M^N(k+1)\rangle_t - \langle M(k), M(k+1)\rangle_t| < \epsilon\right)
$$

$$
= \lim_{N\to\infty} \mathbb{P}\left(\sup_{t\leq T}\left|\int_0^t\left[\sum_r\left(\frac{1}{rNR_{\max}}Y_s^N(r,k+1) + \left(\frac{M}{N} - \sum_{n'}\sum_{r'} n'Y_s^N(r',n')\right)Y_s^N(r,k)\right)\right.\right.\right.
$$

$$
\left.\left.\left. - \left(\int_0^1 \frac{\Lambda}{r}dy_s(r,k+1) + \int_0^1\left(\gamma - \sum_n\int_0^1 n\,dy_s(r,n)\right)dy_s(r,k)\right)\right]ds\right| < \epsilon\right)
$$

$$
\leq \lim_{N\to\infty} \mathbb{P}\left(\sup_{t\leq T} T\left|\sum_r \frac{1}{rNR_{\max}}Y_t^N(r,k+1) - \int_0^1 \frac{\Lambda}{r}dY_t(r,k+1)\right| < \epsilon/3\right)
$$

$$
+ \lim_{N\to\infty} \mathbb{P}\left(\sup_{t\leq T} T\left|\left(\frac{M}{N} - \sum_{n',r'} n'Y_t^N(r',n')\right)Y_t^N(r,k)\right.\right.
$$

$$- \int_0^1 \left(\gamma - \sum_n \int_0^1 n dY_t(r,n) \right) dY_t(r,k) \Bigg| < \epsilon/3 \right)$$

$$+ \lim_{N \to \infty} \mathbb{P} \left(\sup_{t \le T} 2LT|Y_t^N - y_t| < \epsilon/3 \right)$$

$$\le \lim_{N \to \infty} \mathbb{P} \left(\sup_{t \le T} T \left| \frac{C}{\Lambda N R_{\max}} - \frac{\Lambda}{\Lambda/C} \right| \sum_r Y_s(r,k+1) < \epsilon/3 \right)$$

$$+ \lim_{N \to \infty} \mathbb{P} \left(\sup_{t \le T} T \left| \frac{M}{N} - \gamma \right| \sum_r Y_s(r,k) < \epsilon/3 \right)$$

$$\le \lim_{N \to \infty} \mathbb{P} \left(\sup_{t \le T} \frac{CT}{\Lambda} \left| \frac{1}{N R_{\max}} - \Lambda \right| < \epsilon/3 \right) + \lim_{N \to \infty} \mathbb{P} \left(\sup_{t \le T} T \left| \frac{M}{N} - \gamma \right| < \epsilon/3 \right)$$

$$= 0, \tag{A.56}$$

which implies

$$\sup_{t \le T} |\langle \sqrt{N} M^N(k), \sqrt{N} M^N(k+1) \rangle_t - \langle M(k), M(k+1) \rangle_t| \xrightarrow{p} 0. \tag{A.57}$$

Since for all $|i - j| < 1$,

$$\langle \sqrt{N} M^N(i), \sqrt{N} M^N(j) \rangle_t = \langle M(i), M(j) \rangle_t = 0.$$

We can conclude that

$$\sup_{t \le T} |\langle \sqrt{N} M^N(i), \sqrt{N} M^N(j) \rangle_t - \langle M(i), M(j) \rangle_t| \xrightarrow{p} 0. \tag{A.58}$$

for all $0 \le i, j \le K$.

We also know that the jump size of Y_t^N is $1/N$, therefore

$$\lim_{N \to \infty} \mathbb{E} \left[\sup_{0 > t \le T} |\sqrt{N} M_t^N - \sqrt{N} M_{t-}^N| \right] = \lim_{N \to \infty} \mathbb{E} \left[\sup_{0 > t \le T} |\sqrt{N} Y_t^N - \sqrt{N} Y_{t-}^N| \right] = 0. \tag{A.59}$$

By Theorem 1.4 in Chapter 7 of Ethier and Kurtz [8], $\sqrt{N} M_t^N$ converges to the continuous 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. This automatically implies the tightness condition (ii).

For the second term $\int_v^u \sqrt{N}(\beta(Y_z^N) - b(Y_z^N)) dz$, we have by Proposition 5.3 that the quantity $\sqrt{N}(\beta(Y_z^N) - b(Y_z^N))$ 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

$$\sup_{z \in [0,T]} \sqrt{N} |\beta(Y_z^N) - b(Y_z^N)| \le C_1. \tag{A.60}$$

Then,

$$\lim_{\delta \to 0} \lim_{N \to \infty} \mathbb{P} \left(\sup_{u,v \in [0,T], |u-v| \le \delta} \int_v^u \sqrt{N} |\beta(Y_z^N) - b(Y_z^N)| dz < \epsilon \right)$$

$$\le \lim_{\delta \to 0} \lim_{N \to \infty} \mathbb{P} \left(\delta \sup_{z \in [0,T]} \sqrt{N} |\beta(Y_z^N) - b(Y_z^N)| < \epsilon \right)$$

$$\le \lim_{\delta \to 0} \mathbb{P}(\delta C_1 < \epsilon)$$

$$= 0. \tag{A.61}$$

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

Finally, *for the third term*, we have that

$$
\begin{aligned}
\int_v^u \sqrt{N} |b(Y_z^N) - b(y_z)| dz &\leq \int_v^u \sqrt{N} L |Y_z^N - y_z| dz \\
&= \int_v^u L \cdot |D_z^N| dz \\
&\leq L\delta \sup_{t \in [0,T]} |D_t^N|.
\end{aligned}
\tag{A.62}
$$

By Lemma 5.4,

$$
\begin{aligned}
\lim_{\delta \to 0} \lim_{N \to \infty} \mathbb{P} &\left(\sup_{u,v \in [0,T], |u-v| \leq \delta} \int_v^u \sqrt{N} |b(Y_z^N) - b(y_z)| dz < \epsilon \right) \\
&\leq \lim_{\delta \to 0} \lim_{N \to \infty} \mathbb{P} \left(L\delta \sup_{t \in [0,T]} |D_t^N| < \epsilon \right) \\
&\leq \lim_{\delta \to 0} \lim_{N \to \infty} \frac{\mathbb{E}(\sup_{t \in [0,T]} |D_t^N|^2)}{(\epsilon/L\delta)^2} \\
&\leq \lim_{\delta \to 0} \frac{C_0 (L\delta)^2}{\epsilon^2} \\
&= 0,
\end{aligned}
\tag{A.63}
$$

which implies that the oscillation bound holds for the third term. ∎

PROOF: To prove the existence and uniqueness of the SDE (5.3), we show the following two conditions hold: There exists a constant $H < 0$ such that

1. Lipschitz condition: for any $D, \tilde{D} \in \mathbb{R}^{K+1}$, any $t \in (t_0, T)$,

$$
|b'(y_t)D - b'(y_t)\tilde{D}| \leq H|D - \tilde{D}|.
\tag{A.64}
$$

2. Linear growth condition: for any $D \in \mathbb{R}^{K+1}$, any $t \in (t_0, T)$,

$$
|b_+(y_t) + b_-(y_t)| \leq H(1 + |D|), \quad |b'(y_t)D| \leq H(1 + |D|).
\tag{A.65}
$$

By Proposition 4.3, $b(y)$ is Lipschitz, and by our assumption in Theorem 5.1, $b(y)$ is also continuously differentiable, thus $b'(y)$ is bounded. Then, conditions (1) and (2) follow, which prove that there exist a unique solution to the SDE (5.3).

Take expectation on both sides of Eq. (5.9), since $\mathbb{E}[\int_0^t e^{\int_s^t b'(y_u)du} dM_s] = 0$, we have

$$
\mathbb{E}[D_t] = e^{\int_0^t \mathcal{A}(s)ds} \mathbb{E}[D_0].
$$

Therefore,

$$
\begin{aligned}
D_t - \mathbb{E}[D_t] &= e^{\int_0^t \mathcal{A}(s)ds} (D_0 - \mathbb{E}[D_0]) + \int_0^t e^{\int_s^t \mathcal{A}(u)du} dM_s, \\
\Sigma(t) &= E[(D_t - \mathbb{E}[D_t])(D_t - \mathbb{E}[D_t])^\top] \\
&= e^{\int_0^t \mathcal{A}(s)ds} E[(D_0 - \mathbb{E}[D_0])(D_0 - \mathbb{E}[D_0])^\top] (e^{\int_0^t \mathcal{A}(s)ds})^\top \\
&\quad + \left(\int_0^t e^{\int_s^t \mathcal{A}(u)du} dM_s \right) \left(\int_0^t e^{\int_s^t \mathcal{A}(u)du} dM_s \right)^\top
\end{aligned}
\tag{A.66}
$$

$$= e^{\int_0^t \mathcal{A}(s)ds}\Sigma(0)e^{\int_0^t \mathcal{A}^\top(s)ds} + \int_0^t e^{\int_s^t \mathcal{A}(u)du}\mathcal{B}(s)e^{\int_s^t \mathcal{A}^\top(u)du}ds. \tag{A.67}$$

∎

PROPOSITION 7: *For any $s \geq 0$,*

$$\limsup_{N\to\infty} \sqrt{N}|\beta(Y_s^N) - b(Y_s^N)| > \infty. \tag{A.68}$$

PROOF:

$$|\beta(Y_s) - b(Y_s)| \leq 2\sum_n \left|\sum_{r,p,k} \frac{pP_{\max}^N}{rR_{\max}^N}Y_s(r,n,p,k) - \sum_{k\in\mathcal{K}}\iint_{[0,1]\times[0,1]} \frac{p\Lambda}{r}dY_s(r,n,p,k)\right|$$

$$+ 2\sum_n \left|\sum_{r,p,k} pNP_{\max}^N\frac{M}{N}Y_s(r,n,p,k) - \sum_{k\in\mathcal{K}}\iint_{[0,1]\times[0,1]} p\mathcal{P}\gamma dY_s(r,n,p,k)\right|$$

$$+ 2\sum_n \left|\sum_{r,p,k}\left(\sum_{n'}\sum_{r',p',k'} n'Y_s(r',n',p',k')\right)Y_s(r,n,p,k)\right.$$

$$\left. - \sum_{k\in\mathcal{K}}\iint_{[0,1]\times[0,1]}\left(\sum_{n'}\sum_{k\in\mathcal{K}}\iint_{[0,1]\times[0,1]} n'dY_s(r',n',p',k')\right)dY_s(r,n,p,k)\right|.$$

By Eqs. (A.35) and (A.36),

$$|\beta(Y_s) - b(Y_s)| \leq 2(K_{\max}+1)\left(\left|\frac{P_{\max}}{R_{\max}} - \Lambda\right|\frac{C}{\Lambda} + \left|NP_{\max}\frac{M}{N} - \mathcal{P}\gamma\right|\right). \tag{A.69}$$

By the assumptions in Theorem 6.5, we have

$$\limsup_{N\to\infty} \sqrt{N}\left(\frac{P_{\max}^N}{R_{\max}^N} - \Lambda\right) > \infty,$$

$$\limsup_{N\to\infty} \sqrt{N}\left(\frac{M}{N} - \gamma\right) > \infty,$$

$$\limsup_{N\to\infty} \sqrt{N}(NP_{\max}^N - \mathcal{P}) > \infty.$$

Thus,

$$\limsup_{N\to\infty} \sqrt{N}|\beta(Y_s^N) - b(Y_s^N)|$$

$$\leq \limsup_{N\to\infty} 2(K_{\max}+1)\sqrt{N}\left(\left|\frac{P_{\max}}{R_{\max}} - \Lambda\right|\frac{C}{\Lambda} + \left|NP_{\max}\frac{M}{N} - \mathcal{P}\gamma\right|\right)$$

$$> \infty. \tag{A.70}$$

∎

PROPOSITION 8: *For $k > K_{\max}$,*

$$\lim_{N\to\infty} \mathbb{P}\left(\sup_{t\leq T}|\langle\sqrt{N}M^N(k),\sqrt{N}M^N(k+1)\rangle_t - \langle M(k),M(k+1)\rangle_t| < \epsilon\right) = 0. \tag{A.71}$$

PROOF: For the adjacent terms, we have

$$
\lim_{N\to\infty} \mathbb{P}\left(\sup_{t\le T}|\langle\sqrt{N}M^N(k),\sqrt{N}M^N(k+1)\rangle_t - \langle M(k),M(k+1)\rangle_t| < \epsilon\right)
$$

$$
= \lim_{N\to\infty}\mathbb{P}\left(\sup_{t\le T}\left|\int_0^t\left[\sum_{r,p,K\in\mathcal{K}}\left(\frac{pP_{\max}}{rR_{\max}}Y_s^N(r,k+1,p,K)\right.\right.\right.\right.
$$

$$
+ \left(\frac{pNP_{\max}M}{N} - \sum_{n'}\sum_{r',p',K'}n'Y_s^N(r',n',p',K')\right)Y_s^N(r,k,p,K)\right)
$$

$$
- \left(\sum_{K\in\mathcal{K}}\iint_{[0,1]\times[0,1]}\frac{p\Lambda}{r}dY_s(r,k+1,p,K)\right.
$$

$$
+ \left.\left.\left.\sum_{K\in\mathcal{K}}\iint_{]0,1]\times[0,1]}p\mathcal{P}\left(\gamma - \sum_n\sum_{K\in\mathcal{K}}\iint_{]0,1]\times[0,1]}ndy_s(r,n,p,K)\right)dy_s(r,k,p,K)\right)\right]ds\right| < \epsilon\right)
$$

$$
\le \lim_{N\to\infty}\mathbb{P}\left(\sup_{t\le T}T\left|\sum_{r,p,K}\frac{pP_{\max}}{rR_{\max}}Y_t^N(r,k+1,P,K) - \sum_{K\in\mathcal{K}}\iint_{[0,1]\times[0,1]}\frac{p\Lambda}{r}dY_s(r,k+1,p,K)\right| < \epsilon/3\right)
$$

$$
+ \lim_{N\to\infty}\mathbb{P}\left(\sup_{t\le T}T\left|\sum_{r,p,K\in\mathcal{K}}\left(pNP_{\max}\frac{M}{N} - \sum_{n'}\sum_{r',p',K'}n'Y_s^N(r',n',p',K')\right)Y_t^N(r,k,p,K)\right.\right.
$$

$$
- \left.\left.\sum_{K\in\mathcal{K}}\iint_{]0,1]\times[0,1]}p\mathcal{P}\left(\gamma - \sum_n\sum_{K\in\mathcal{K}}\iint_{]0,1]\times[0,1]}ndy_s(r,n,p,K)\right)dy_s(r,k,p,K)\right| < \epsilon/3\right)
$$

$$
+ \lim_{N\to\infty}\mathbb{P}\left(\sup_{t\le T}2LT\left|Y_t^N - y_t\right| < \epsilon/3\right)
$$

$$
\le \lim_{N\to\infty}\mathbb{P}\left(\sup_{t\le T}T\left|\frac{CP_{\max}}{\Lambda R_{\max}} - \frac{\Lambda}{\Lambda/C}\right|\sum_{r,p,K}Y_s(r,k+1,p,K) < \epsilon/3\right)
$$

$$
+ \lim_{N\to\infty}\mathbb{P}\left(\sup_{t\le T}T\left|NP_{\max}\frac{M}{N} - \mathcal{P}\gamma\right|\sum_{r,p,K}pY_s(r,k,p,K) < \epsilon/3\right)
$$

$$
\le \lim_{N\to\infty}\mathbb{P}\left(\sup_{t\le T}\frac{CT}{\Lambda}\left|\frac{P_{\max}}{R_{\max}} - \Lambda\right| < \epsilon/3\right) + \lim_{N\to\infty}\mathbb{P}\left(\sup_{t\le T}T\left|NP_{\max}\frac{M}{N} - \mathcal{P}\gamma\right| < \epsilon/3\right)
$$

$$
= 0, \tag{A.72}
$$

which implies

$$
\sup_{t\le T}|\langle\sqrt{N}M^N(k),\sqrt{N}M^N(k+1)\rangle_t - \langle M(k),M(k+1)\rangle_t| \xrightarrow{p} 0. \tag{A.73}
$$

∎