Rate Control under Heavy Traffic with Strategic Servers

Erhan Bayraktar* Amarjit Budhiraja† Asaf Cohen

February 21, 2017

Abstract

We consider a large queueing system that consists of many strategic servers that are weakly
interacting. Each server processes jobs from its unique critically loaded buffer and controls
the rate of arrivals and departures associated with its queue to minimize its expected cost.
The rates and the cost functions in addition to depending on the control action, can depend,
in a symmetric fashion, on the size of the individual queue and the empirical measure
of the states of all queues in the system. In order to determine an approximate Nash
equilibrium for this finite player game we construct a Lasry-Lions type mean-field game
(MFG) for certain reflected diffusions that governs the limiting behavior. Under conditions,
we establish the convergence of the Nash-equilibrium value for the finite size queuing system
to the value of the MFG. In general closed form solutions of such MFG are not available
and thus numerical methods are needed. We use the Markov chain approximation method
to construct approximations for the solution of the MFG and establish convergence of the
numerical scheme over a small time interval.

AMS Classification: 60K25, 91A13, 60K35, 93E20, 60H30, 65M12, 60F17.

Keywords: Heavy traffic limits, queuing systems, strategic servers, mean-field games, rate
control, reflected diffusions.

1 Introduction

Rate controlled queueing systems commonly arise from applications in communication sys-
tems, see e.g. [26] and references therein. Recently, they have also been considered
in modeling limit order books, see e.g. [8, 9, 36]. A common approach to the
study of such rate control problems when the system is in heavy traffic is through diffusion
approximations. In a problem setting where there is interaction between servers/queues in
that the rates or costs associated with a particular queue and server can depend on the states
of the other queues, this approach leads to a stochastic control problem for n-dimensional
reflected diffusions, where n is the number of queues in the system. When n is large such con-
trol problems are computationally intractable and in general this ‘curse of dimensionality’ is

*Research partially supported by the National Science Foundation (DMS-1613170) and the Susan M. Smith
Professorship.
†Research partially supported by the National Science Foundation (DMS-1305120), the Army Research Office
(W911NF-14-1-0331) and DARPA (W911NF-15-2-0122).
unavoidable. However, when there are certain symmetries present and the interaction between queues is weak, in that each queue has $O(1/n)$ affect on any other queue in the system, a natural approach is to consider, in addition to heavy traffic, another asymptotic regime where the number of queues $n$ approaches $\infty$ as well. Such model settings arise in many applications, e.g., cloud computing, live streaming, limit order books, customer service systems, etc. In many of these contexts the servers are strategic, for example, in customer service networks, servers respond to workload incentives (see [32]), and in the context of limit order books buyers and sellers place their orders in a strategic manner and interact weakly through their impact on the price distribution.

We study one of the simplest forms of queuing networks, namely a collection of $n$ single server queues. The system is assumed to be critically loaded. For this, we fix an arbitrary sequence of scaling parameters $\{e_n\}_n$ that satisfy $e_n \to \infty$ when $n \to \infty$. We assume that, for each queue, the traffic intensity, which is the ratio between the rate of arrivals to the rate of service is $1 - O(1/\sqrt{e_n})$ (see [23]). The arrival and service rates for each queue are $O(e_n)$. In absence of control and state dependence of rates, the analysis simply reduces to that of a single $M/M/1$ queue which in the heavy traffic can be approximated by a one dimensional reflected Brownian motion. In the setting we consider, every server can exercise control on the arrival and service rates associated with its own queue. This control which is $O(\sqrt{e_n})$ is of lower order compared to the overall rate but it can have significant impact on performance in the asymptotic regime we consider. In the heavy traffic regime with a fixed number of servers, performance improvement using an $O(\sqrt{e_n})$ control has been well studied (see for example [44, Chapter 9]). In a setting (that is quite different from the one considered here) where the number of nodes/servers approach $\infty$, numerical results that show performance improvement under $O(\sqrt{e_n})$ controls can be found in [18]. In the model we consider, the rates can depend on the state of the individual queue, furthermore a particular queue’s state is influenced by the remaining queue states through their empirical measure. The control action for each server can use information on the history of queue lengths, arrival and processing times and control actions for all the queues in the system. We consider a rate control problem where each server aims to minimize its individual cost. Although many different types of cost criteria can be considered, here for simplicity we consider a cost function over a finite time horizon. This cost function may depend on the individual queue lengths, the control action, and the empirical measure of all the queue states. A natural way to formulate optimality for such $n$-player games is through the notion of a (near) Nash equilibrium. Computing Nash equilibria, even if one considers the simplified heavy traffic approximation in terms of reflected diffusions, is computationally a very challenging problem when $n$ is large. The goal of this work is to analyze an asymptotic formulation where simultaneously each queue approaches heavy traffic and the number of queues approaches $\infty$. There is extensive literature on heavy traffic limits of rate controlled queuing systems with a fixed number of queues [44, 4, 19], however to the best of our knowledge this is the first work to study the asymptotics for controlled queues where in addition to heavy traffic, the number of queues approach infinity as well. We will show that the asymptotics in this regime is governed by a Lasry-Lions type mean field game (MFG) for reflected diffusions. We will use the solution of the MFG to construct an asymptotic Nash equilibrium for the $n$-player queuing system. This equilibrium has a simple and appealing decentralized structure: each server bases his decision only upon the length of his own queue.
and a deterministic measure valued function obtained from the solution of the MFG. We also prove that the value of the above near Nash equilibrium converges to the value function of the MFG as \( n \to \infty \). In general closed form solutions for MFG of the form that arise here are not available and thus one needs numerical approximations. In our work we study one such procedure that uses the Markov chain approximation method ([45]) and establish convergence of the scheme over a small time interval.

The theory of mean field games was initiated a decade ago in the seminal work of Lasry and Lions [49, 50, 51], and Huang, Malhamé, and Caines [39, 38]. In recent years there has been a growing interest in this field. For recent theoretical developments and applications of this theory see [20, 35, 31, 21, 18, 22, 47, 27] and references therein. Mean field approximations for weakly interacting stochastic particles have a long history starting from the works of Boltzmann, McKean, Kac and others (see [57] and references therein). Even in the context of queuing systems and communication networks, there have been many works [30, 5, 40, 60, 16, 33, 11, 14, 34, 3, 18]. Another related branch is agent based models with mean-field interaction (but without strategic agents), see e.g. [29, 6, 7]. MFGs have also been used by queuing theorists as a tool in recent years see e.g. [51, 61, 52, 46]. In contrast to these works we consider an MFG associated with queuing systems in heavy traffic.

We now make some comments on proof techniques. Roughly speaking, solution of an MFG considered here can be viewed as the solution of a fixed point problem on the space of probability measures on certain path spaces (see Section 3). In order to characterize such solutions, in a setting where each agent’s state evolution is described through a diffusion process, Lasry and Lions [49, 50, 51] studied wellposedness of two coupled nonlinear partial differential equations; one is an equation of Hamilton-Jacobi-Bellman (HJB) type while the second takes the form of a Kolmogorov forward equation. A somewhat different approach is taken in recent works of Carmona, Delarue, and Lacker [21, 22]. Using probabilistic methods, authors characterize the MFG solution as a solution to certain forward backward stochastic differential equations. In all the above papers the \( n \)-player system is described through a collection of stochastic differential equations and the mean field game gives the asymptotic behavior of the system as \( n \to \infty \). In contrast, in the current work, for a fixed value of \( n \) the state process is given through a collection of controlled jump-Markov processes with jump sizes that approach 0 as \( n \to \infty \). Thus here we have two forms of asymptotic behavior, one corresponds to the large agent limit while the other corresponds to diffusion approximations of pure jump processes with vanishing jump sizes. Analyzing this simultaneous limit behavior, which is key to the proof of Theorem 4.1 that identifies an asymptotic Nash equilibrium for the \( n \)-player game as \( n \to \infty \), requires new techniques. A key ingredient is to prove suitable tightness properties and to analyze and characterize the weak limit points of certain stochastic processes and random measures. Here we make use of a result from [12] (see Lemma 4.2 therein) which can be viewed as an extension of de Finetti’s theorem for sequences of random measures on certain path spaces. In proving suitable tightness properties of control processes we consider a relaxed control formulation. Our assumption on the uniqueness of the minimizer (Assumption 3.1(b)) ensures that extending the class of controls in this manner does not lower the cost. Such relaxed control formulations have been widely used in stochastic control theory, see e.g. [28, 23, 15, 45]. More recently, they have also been invoked in the study of mean field games [48, 47, 27]. Other key ingredients
needed for the proof are (a) regularity properties of the optimal feedback controls for certain stochastic control problems for reflected diffusions; (b) weak formulation of stochastic control problems for reflected diffusions. We find that the approach based on HJB characterizations of value functions is particularly well suited for our problem setting. In particular, regularity results from [53] for quasilinear PDE with Neumann boundary conditions give us the required estimates for obtaining the desired properties of the optimal feedback controls. Use of HJB theory in our analysis is a feature that is common to the approach initiated by Lasry and Lions, however a point of departure is that instead of using Kolmogorov forward equations we characterize probability distributions as unique weak solutions of suitable reflected diffusions. In this sense our approach is closer to [22] although, unlike [22], we do not make use of forward backward stochastic differential equations in our work. In the final section of this work we provide a numerical scheme for solving the MFG and under conditions prove convergence over a small time interval (Theorem 5.1). Viscosity solution methods for numerical approximations of MFG have been studied in [1]. Here we consider a probabilistic approach based on a Markov chain approximation method [45]. The convergence proof once more relies on weak convergence arguments of the form used in the proof of Theorem 4.1. The key result is Proposition 5.1 which establishes an ‘almost contraction’ property that is uniform in the discretization parameter over a fixed small interval. To the best of our knowledge this is the first work to develop a numerical scheme based on Markov chain approximations for an MFG problem.

In summary our main contributions are as follows. We

- consider a rate control problem for large symmetric queuing systems in heavy traffic with strategic servers;
- introduce an MFG for controlled reflected diffusions, and in Theorem 3.1 establish its solvability under Assumption 3.1 and prove unique solvability assuming in addition Assumption 3.2;
- use the solution of a diffusion MFG to construct, under Assumptions 3.1 and 4.1 an asymptotically optimal Nash equilibrium for the \( n \) player countable state game using techniques that combine heavy traffic analysis with large agent asymptotics (Theorem 4.1);
- develop a numerical scheme to solve the MFG that under Assumptions 3.1, 5.1, and 5.2 converges over a small time interval (Theorem 5.1).

Thus Assumption 3.1 is a basic assumption for all our results. This assumption imposes Lipschitz continuity of the various functions in the model. Assumption 3.2 that is introduced for the uniqueness of the solution of the MFG, is common in the literature of MFG. It says that the drift is independent of the mean-field term and that the running cost and terminal cost satisfy a certain monotonicity property, see Remark 3.1 for a discussion of the assumption. In order to obtain a asymptotic Nash-equilibrium, we require, in addition to Assumption 3.1 Assumption 4.1 which as the first part of Assumption 4.2 says that the drift is free of the mean-field term and in addition requires a basic convergence property for the initial conditions . For this result we do not require the monotonicity property, in particular the uniqueness of the solution of the MFG is not needed. For the convergence of the numerical scheme, we require, in addition to Assumption 3.1 Assumption 5.1 that imposes some regularity conditions on the
We use the following notation. For some auxiliary results from Section 5. We provide a numerical scheme for solving the MFG. Section 6 provides the result on its solvability (Theorem 3.1). Section 4 constructs an asymptotic Nash equilibrium for the n-player game and the main result of this section is Theorem 4.1 which proves asymptotic optimality. In Section 5 we provide a numerical scheme for solving the MFG. Section 6 provides proofs of some auxiliary results from Section 5.

1.1 Notation

We use the following notation. For $t \in (0, \infty)$ and a function $f : [0, \infty) \to \mathbb{R}^d$, $\|f\|_t = \sup_{0 \leq s \leq t} \|f\|$. For metric spaces $S_1, S_2$ denote by $C(S_1 : S_2)$ the space of $S_2$ valued continuous functions on $S_1$. When $S_2 = \mathbb{R}$, we abbreviate the notation to simply $C(S_1)$. The space $C([0, T] : S)$ for a Polish space $S$ will be equipped with the uniform topology. We will denote by $D([0, T] : S)$ the space of $S$ valued functions that are right continuous and have left limits (RCLL) defined on $[0, T]$. This space is equipped with the usual Skorohod topology. For $f \in D([0, T] : \mathbb{R}^d)$ and $0 \leq t \leq t' \leq T$, $\|f\|_t = \sup_{0 \leq s \leq t} \|f(s)\|$. In case that $d = 1$, we often use $|f|_t$. We will denote by $\text{Lip}_1(S)$ the space of real Lipschitz functions on $S$ whose Lipschitz norm is bounded by 1, namely the class of functions $f : S \to \mathbb{R}$ with

$$|f(x) - f(y)| \leq d(x, y), \quad x, y \in S$$

where $d$ denotes the metric on $S$. Denote by $P(S)$ the space of probability measures on $S$. We endow $P(S)$ with the topology of weak convergence of measures. Convergence in distribution of $S$ valued random variable $X_n$ to $X$ will be denoted as $X_n \Rightarrow X$. For $T, L \in (0, \infty)$, the space $P(C([0, T] : [0, L]))$ will be denoted as $P_{T, L}$. The Wasserstein distance of order 1 on $P(S)$, where $S$ is a compact metric space, is defined as

$$W_1(\eta', \eta) = \inf \left\{ \left[ \int_S d(x, y) d\pi(x, y) \right] : \pi \in P(S \times S) \text{ with marginals } \eta' \text{ and } \eta \right\},$$

where $\eta, \eta' \in P(S)$. We denote by $C^{1, 2}([0, T] \times [0, L])$ the space of functions $f : (0, T) \times (0, L) \to \mathbb{R}$ that are continuously differentiable (resp., twice continuously differentiable) with respect to (w.r.t.) the first (resp., second) variable and are such that the function and the derivatives extend continuously to $[0, T] \times [0, L]$. For $\phi \in C^{1, 2}([0, T] \times [0, L])$, $D_1 \phi, D_2 \phi, D^2 \phi$ will denote the time derivative and the first two space derivatives of $\phi$, respectively. For $x \in S$, $\delta_x \in P(S)$ denotes the Dirac measure at $x$.

2 The n-server queueing control problem

We consider a symmetric $n$-server stochastic queueing system. Each server $i \in \{1, \ldots, n\}$ is associated with a queue with a finite capacity and is able to control the rates of the service and arrivals of jobs to its queue. The rates can also depend on the state of the queue and on the empirical measure of the states of all the $n$-queues. We will consider a regime where
the arrival and processing rates are approximately the same and are of the same order as the number of queues in the system. We now describe our precise scaling regime and introduce the various processes that determine the evolution of the state of the system.

2.1 Diffusion scaling

Fix $T, L > 0$ and $n \in \mathbb{N}$. Here $T$ denotes the terminal time of our finite time horizon and $[0, L]$ will be the state space of the scaled queue length process. Let $U$ be a compact subset of $\mathbb{R}$. We will denote by $\lambda^{n,i}$ and $\mu^{n,i}$ the (controlled) arrival and service rates associated with queue $i$. Fix an arbitrary sequence of scaling parameters $\{e_n\}_n$ that satisfy $e_n \to \infty$ when $n \to \infty$. The rates will be $O(\sqrt{e_n})$ perturbations of certain nominal (uncontrolled) $O(e_n)$ overall rates. More precisely, we assume that there exist $\hat{\lambda}, \hat{\mu} > 0$, and bounded and measurable functions $\lambda, \mu : [0, T] \times \mathcal{P}([0, L]) \times [0, L] \times U \to \mathbb{R}$ such that

$$
\begin{align*}
\lambda^{n,i}(t) &= \hat{\lambda}_n + \lambda(t, \tilde{\nu}^n(t), \tilde{Q}^{n,i}(t), \alpha^{n,i}(t)) \sqrt{e_n} + o(\sqrt{e_n}), \\
\mu^{n,i}(t) &= \hat{\mu}_n + \mu(t, \tilde{\nu}^n(t), \tilde{Q}^{n,i}(t), \alpha^{n,i}(t)) \sqrt{e_n} + o(\sqrt{e_n}),
\end{align*}
$$

where $\tilde{Q}^{n,i}(t) = \frac{1}{\sqrt{e_n}} Q^{n,i}(t)$, $Q^{n,i}(t)$ is the size of the $i$-th queue at time $t$,

$$
\tilde{\nu}^n(t) = \frac{1}{n} \sum_{i=1}^{n} \delta_{Q^{n,i}(t)}
$$

is the empirical distribution of the scaled queue lengths in the $n$-th system at time $t$, and $\alpha^{n,i}(t)$ is the control that server $i$ exercises at time $t$. The term $o(\sqrt{e_n})$ represents an expression of the form $r_n(t, \tilde{\nu}^n(t), \tilde{Q}^{n,i}(t), \alpha^{n,i}(t))$ where $r_n : [0, T] \times \mathcal{P}([0, L]) \times [0, L] \times U \to \mathbb{R}$ are measurable functions such that $r_n/\sqrt{e_n}$ converges uniformly to 0 as $n \to \infty$. Additional conditions on $\lambda$ and $\mu$ will be introduced in later sections. Each server in the $n$-th system has a finite buffer of size $L_n = \sqrt{e_n} L$ and arriving jobs to a full buffer are lost. We assume that the system is in heavy traffic, namely

$$
\hat{\lambda} = \hat{\mu}.
$$

We now give a precise description of controlled stochastic processes of interest. For each fixed $n$, let $(\Omega', \mathcal{F}', \mathbb{P}')$ be a probability space on which are given unit rate independent Poisson processes $N^{n,i}$ and $M^{n,i}$, $i = 1, \ldots, n$. Roughly speaking, $N^{n,i}$ will correspond to the stream of jobs entering the $i$-the queue and $M^{n,i}$ to the jobs that leave the system after service. The evolution of the $i$-th controlled queue is given as follows.

$$
Q^{n,i}(t) = Q^{n,i}(0) + A^{n,i}(t) - D^{n,i}(t), \quad i = 1, \ldots, n, \quad t \in [0, T],
$$

where $Q^{n,i}(0) \geq 0$ is the initial size of the $i$-th queue,

$$
A^{n,i}(t) = N^{n,i} \left( \int_0^t 1_{\{Q^{n,i}(s) < L\}} \lambda^{n,i}(s) ds \right), \quad D^{n,i}(t) = M^{n,i} \left( \int_0^t 1_{\{Q^{n,i}(s) > 0\}} \mu^{n,i}(s) ds \right)
$$

are the arrival and departure processes respectively, where $\lambda^{n,i}$ and $\mu^{n,i}$ are defined as in (2.1) in terms of control processes $\{\alpha^{n,i}\}$. We will assume that $\{Q^{n,i}(0)\}_{i=1}^{n}$ are exchangeable for all $n$. We require that the control processes are suitably non-anticipative. Specifically, we
assume that for some filtration \( \{ F_t \} \) on \((\Omega',\mathcal{F}',\mathbb{P}')\), \( \alpha^{n,i} \) is \( \{ F_t \} \)-progressively measurable, \( Q^{n,i} \), \( n \geq i \geq 1 \) are \( \{ F_t \} \)-adapted, and

\[
\hat{A}^{n,i}(t) = \frac{A^{n,i}(t) - \int_0^t 1_{\{Q^{n,i}(s) < L\}} \lambda^{n,i}(s) ds}{\sqrt{e_n}}, \quad \hat{D}^{n,i}(t) = \frac{D^{n,i}(t) - \int_0^t 1_{\{Q^{n,i}(s) > 0\}} \mu^{n,i}(s) ds}{\sqrt{e_n}}
\]

are \( \{ F_t \} \) martingales with quadratic variations

\[
\langle \hat{A}^{n,i}, \hat{A}^{n,j} \rangle(t) = \delta_{ij} \frac{1}{e_n} \int_0^t 1_{\{Q^{n,i}(s) < L\}} \lambda^{n,i}(s) ds, \quad \langle \hat{D}^{n,i}, \hat{D}^{n,j} \rangle(t) = \delta_{ij} \frac{1}{e_n} \int_0^t 1_{\{Q^{n,i}(s) > 0\}} \mu^{n,i}(s) ds,
\]

and \( \langle \hat{A}^{n,i}, \hat{D}^{n,j} \rangle(t) = 0, t \in [0,T], i, j = 1, \ldots, n \) where \( \delta_{ij} = 1 \) if \( i = j \) and 0 otherwise. The process \( \alpha^n = \{ \alpha^{n,i} \}_{i=1}^n \) will be referred to as an admissible control and we denote the collection of all such controls by \( \mathcal{U}^n \). With an abuse of terminology, for \( \alpha^n = \{ \alpha^{n,i} \}_{i=1}^n \) as above, we will refer to \( \alpha^n \) as an admissible control for the \( i \)-th player.

From (2.2) we have the following evolution equation for the scaled queue length processes. For \( t \in [0,T] \)

\[
\hat{Q}^{n,i}(t) = \hat{Q}^{n,i}(0) + \hat{A}^{n,i}(t) - \hat{D}^{n,i}(t) + \int_0^t \hat{b}^{n,i}(s) ds + \hat{Y}^{n,i}(t) - \hat{R}^{n,i}(t) + o(1),
\]

where

\[
\hat{Y}^{n,i}(t) = \frac{\int_0^t 1_{\{\hat{Q}^{n,i}(s) = 0\}} \mu^{n,i}(s) ds}{\sqrt{e_n}}, \quad \hat{R}^{n,i}(t) = \frac{\int_0^t 1_{\{\hat{Q}^{n,i}(s) = L\}} \lambda^{n,i}(s) ds}{\sqrt{e_n}},
\]

\[
\hat{b}^{n,i}(t) = b(t, \hat{v}^{n,i}(t), \hat{Q}^{n,i}(t), \alpha^{n,i}(t)), \quad b = \lambda - \mu,
\]

and \( o(1) \) represents a stochastic process \( \eta^{n,i} \) satisfying for each \( i \) sup\(_{0 \leq t \leq T} |\eta^{n,i}(t)| \to 0 \), in probability, as \( n \to \infty \). The above dynamics can equivalently be described in terms of a Skorohod map as we do below. Let \( \mathcal{D}_0[0,T] \) be the subset of \( \mathcal{D}([0,T] : \mathbb{R}) \) consisting of all \( \psi \) such that \( \psi(0) \in [0,L] \).

**Definition 2.1** Given \( \psi \in \mathcal{D}_0[0,T] \) we say the triplet of functions \( (\varphi, \zeta_1, \zeta_2) \in \mathcal{D}([0,T] : \mathbb{R}^3) \) solve the Skorohod problem for \( \psi \) if the following properties are satisfied:

(i) For every \( t \in [0,T] \), \( \varphi(t) = \psi(t) + \zeta_1(t) - \zeta_2(t) \in [0,L] \).

(ii) \( \zeta_i \) are nonnegative and nondecreasing, \( \zeta_1(0) = \zeta_2(0) = 0 \), and

\[
\int_{[0,L]} 1_{[0,L]}(\varphi(s)) d\zeta_1(s) = \int_{[0,T]} 1_{[0,L]}(\varphi(s)) d\zeta_2(s) = 0.
\]

We denote by \( \Gamma(\psi) = (\Gamma_1, \Gamma_2, \Gamma_3)(\psi) = (\varphi, \zeta_1, \zeta_2) \) and refer to \( \Gamma \) as the Skorohod map.

It is known that there is a unique solution to the Skorohod problem for every \( \psi \in \mathcal{D}_0[0,T] \) and so the Skorohod map in Definition 2.1 is well defined. The Skorohod map has the following Lipschitz property (see [13]).
Lemma 2.1 There exists $c_S \in (0, \infty)$ such that for all $\omega, \tilde{\omega} \in \mathcal{D}_0([0, T])$,
\[ \sum_{i=1}^{3} \|F_i(\omega) - F_i(\tilde{\omega})\|_T \leq c_S \|\omega - \tilde{\omega}\|_T. \]

The dynamics in (2.5) can be described in terms of the Skorohod map as follows
\[ (Q^{n,i}, \tilde{Y}^{n,i}, \tilde{R}^{n,i})(t) = \Gamma \left( \tilde{Q}^{n,i}(0) + \int_{0}^{t} \tilde{b}^{n,i}(s) ds + \tilde{A}^{n,i}() - \tilde{D}^{n,i}() + o(1) \right)(t), \quad t \in [0, T]. \]  

(2.7)

2.2 The control problem

We now introduce the cost function and the control problem studied in this work. The total expected cost for server $i$ associated with the initial condition $Q^n(0) = (Q^{n,1}(0), \ldots, Q^{n,n}(0))$ and control $\alpha^n = (\alpha^{n,1}, \ldots, \alpha^{n,n}) \in \mathcal{U}^n$ is given by
\[
J^{n,i}(Q^n(0); \alpha^n) \doteq \mathbb{E} \left[ \int_{0}^{T} f(t, \tilde{v}^{n}(t), Q^{n,i}(t), \alpha^{n,i}(t)) dt + g(\tilde{v}^{n}(T), Q^{n,i}(T)) \right. \\
\left. + \int_{0}^{T} y(t, \tilde{v}^{n}(t)) d\tilde{Y}^{n,i}(t) + \int_{0}^{T} r(t, \tilde{v}^{n}(t)) d\tilde{R}^{n,i}(t) \right].
\]

where $f : [0, T] \times \mathcal{P}([0, L]) \times [0, L] \times U \to \mathbb{R}$ is the running cost, $g : \mathcal{P}(\mathbb{R}) \times [0, L] \to \mathbb{R}$ is the terminal cost, and $r, y : [0, T] \times \mathcal{P}([0, L]) \to \mathbb{R}_+$ are the costs associated with rejection of jobs and empty buffers, respectively. Here $f, g, y,$ and $r$ are bounded measurable functions that will be required to satisfy additional conditions that will be introduced in Section 3 (see Assumption 3.1). Each player (server) seeks to minimize its cost. A natural formulation of optimality for such $n$-player games is given in terms of a (near) Nash equilibrium. Computing near Nash equilibria for such complex and large multi-player games is in general intractable and thus we instead consider an asymptotic formulation of the problem.

Definition 2.2 A sequence of admissible controls $\{\tilde{\alpha}^{n,i} : 1 \leq i \leq n\}_{n \in \mathbb{N}}$ is called an asymptotic Nash equilibrium if for every player $j$, and every sequence of admissible controls $\{\beta^n\}_{n=1}^\infty$ for that player, one has
\[
\lim_{n \to \infty} \sup J^{n,j}(Q^n(0); \tilde{\alpha}^{n,1}, \ldots, \tilde{\alpha}^{n,n}) \leq \lim_{n \to \infty} \inf J^{n,j}(Q^n(0); \tilde{\alpha}^{n,1}, \ldots, \tilde{\alpha}^{n,n}, \beta^{n}, \tilde{\alpha}^{n,j+1}, \ldots, \tilde{\alpha}^{n,n}).
\]

Objective of this work is to show that, under conditions, an asymptotic Nash equilibrium exists which can be numerically computed (at least approximately) by analyzing a related MFG. The main results are Theorem 3.1 (solvability of MFG), Theorem 4.1 (asymptotic optimality), and Theorem 5.1 (convergence of the numerical scheme).

3 The MFG

A natural approach for constructing asymptotic near Nash equilibria for the above $n$-player game has emerged from the works of [20, 21, 22]. Starting point in this approach is to consider an MFG that formally corresponds to the limit of the above $n$-player games as $n \to \infty$. In this section we give a precise description of this MFG in the current context and give our main results on existence and uniqueness of solutions.
3.1 Description of the MFG

The basic idea in the formulation of the MFG is to approximate the scaled queue length process for a typical queue by a suitable drift-controlled reflected Brownian motion. We next introduce this controlled process. Let \((\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})\) be a filtered probability space on which is given a one dimensional standard \(\mathcal{F}_t\)-Brownian motion \(B\). We will refer to the collection \((\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P}, B)\) as a system and denote it by \(\Xi\). Given \(x \in [0, L]\), \(t \in [0, T]\), and \(\nu \in \mathcal{P}_{T,L}\), we denote by \(\mathcal{A}(\Xi, t, x, \nu)\) the collection of all pairs \((\alpha, Z)\) where \(\alpha = \{\alpha(s)\}_{0 \leq s \leq T-t}\) is a \(U\)-valued \(\mathcal{F}_s\)-progressively measurable process, \(Z = \{Z(s)\}_{0 \leq s \leq T-t}\) is a \([0, L] \times \mathbb{R}_+ \times \mathbb{R}_+\) valued \(\mathcal{F}_s\)-adapted continuous process such that, \(Z = (X, Y, R)\) and

\[
Z(s) = (X, Y, R)(s) = \Gamma \left( x + \int_0^s b(u)du + \sigma B(\cdot) \right)(s), \quad s \in [0, T-t],
\]

where

\[
b(u) \doteq b(t + u, \nu(t + u), X(u), \alpha(u)), \quad u \in [0, T-t],
\]

\(\nu(s)\) is the marginal of \(\nu\) at time instant \(s\) and \(\sigma = \sqrt{2\lambda}\). We now introduce the cost function in the MFG. Given \(\nu \in \mathcal{P}_{T,L}, x \in [0, L], t \in [0, T]\), and a system \(\Xi\) as above, let \((\alpha, Z) \in \mathcal{A}(\Xi, t, x, \nu)\). Define

\[
J_\nu(t, x, \alpha, Z) \doteq \mathbb{E}\left[ \int_0^{T-t} f(s + t, \nu(s + t), X(s), \alpha(s))ds + g(\nu(T), X(T - t)) 
+ \int_0^{T-t} y(s + t, \nu(s + t))dY(s) + \int_0^{T-t} r(s + t, \nu(s + t))dR(s) \right].
\]

We define the value function associated with the above cost as:

\[
V_\nu(t, x) = \inf_{(\alpha, Z) \in \mathcal{A}(\Xi, t, x, \nu)} J_\nu(t, x, \alpha, Z).
\]

**Definition 3.1** A solution to the MFG with initial condition \(x \in [0, L]\) is defined to be a \(\nu \in \mathcal{P}_{T,L}\) such that there exist a system \(\Xi\) and an \((\alpha, Z) \in \mathcal{A}(\Xi, 0, x, \nu)\) such that \(Z = (X, Y, R)\) satisfies \(\mathbb{P} \circ X^{-1} = \nu\) and

\[
V_\nu(0, x) = J_\nu(0, x, \alpha, Z).
\]

If there exists a unique such \(\nu\), we refer to \(V_\nu(0, x)\) as the value of the MFG with initial condition \(x\).

To find a solution of the MFG one usually follows the following iterative procedure:

(i) For a fixed \(\nu \in \mathcal{P}_{T,L}\) solve the stochastic control problem (3.3) with \(t = 0\), namely find a system \(\Xi\) and \((\alpha, Z) \in \mathcal{A}(\Xi, 0, x, \nu)\) such that (3.4) holds. Denote by \(\tilde{\nu}\) the law of \(X\) where \(Z = (X, Y, R)\) and write \(\tilde{\nu} = \Phi(\nu)\) (this is not precise since in general there may be more than one solution of the stochastic control problem in (3.3)).

(ii) Find the fixed point of the map \(\Phi\), namely a \(\tilde{\nu} \in \mathcal{P}_{T,L}\) for which \(\tilde{\nu} = \Phi(\tilde{\nu})\). Note that by definition, such a \(\tilde{\nu}\) will be a solution of the MFG.

We now analyze the MFG by following the above steps. The main result is Theorem 3.1 which gives existence of solutions of the MFG under suitable conditions and proves uniqueness of solutions under stronger conditions.
3.2 Solving the stochastic control problem

For \( c \in (0, \infty) \), let \( \mathcal{M}_c \) be the collection of all \( \nu \in \mathcal{P}_{T,L} \) such that

\[
\sup_{0 \leq s < t \leq T} \frac{W_1(\nu(t), \nu(s))}{(t - s)^{1/2}} \leq c
\]

and let

\[
\mathcal{M}_0 = \bigcup_{c > 0} \mathcal{M}_c.
\]

Fix a measure \( \nu \in \mathcal{M}_0 \). For \( \nu \in \mathcal{P}_{T,L} \), the function \( t \mapsto \nu(t) \) is a continuous function from \([0, T]\) to \( \mathcal{P}_{[0, L]} \) and with an abuse of notation we denote this continuous function once more as \( \nu \).

As one might expect, the value function \( V_\nu(t, x) \) corresponds to the Hamilton-Jacobi-Bellman (HJB) equation:

\[
-D_t \phi - H(t, \nu(t), x, D\phi) - \frac{1}{2} \sigma^2 D^2 \phi = 0, \quad (t, x) \in [0, T] \times [0, L],
\]

with the boundary conditions (BC)

\[
\phi(T, x) = g(\nu(T), x), \quad D\phi(t, 0) = -y(t, \nu(t)), \quad \text{and} \quad D\phi(t, L) = r(t, \nu(t)), \quad t \in [0, T],
\]

where \( H \) is the Hamiltonian given as

\[
H(t, \eta, x, p) = \inf_{u \in U} h(t, \eta, x, u, p)
\]

and \( h : [0, T] \times \mathcal{P}([0, L]) \times [0, L] \times U \times \mathbb{R} \to \mathbb{R} \) is defined as

\[
h(t, \eta, x, u, p) = f(t, \eta, x, u) + b(t, \eta, x, u)p.
\]

We now introduce a key condition under which the above HJB equation characterizes the value function \( V_\nu(t, x) \).

**Assumption 3.1**

(a) There exists \( c_L \in (0, \infty) \) such that for every \( t, t' \in [0, T] \), \( \eta, \eta' \in \mathcal{P}([0, L]) \), \( x, x' \in [0, L] \), and \( \alpha, \alpha' \in U \),

\[
|f(t, \eta, x, \alpha) - f(t', \eta', x', \alpha')| + |g(\eta, x) - g(\eta', x')| + |b(t, \eta, x, \alpha) - b(t', \eta', x', \alpha')| \\
+ |y(t, \eta) - y(t', \eta')| + |r(t, \eta) - r(t', \eta')| \\
\leq c_L(|t - t'| + W_1(\eta, \eta') + |x - x'| + |\alpha - \alpha'|).
\]

(b) For every \( (t, \eta, x, p) \in [0, T] \times \mathcal{P}([0, L]) \times [0, L] \times \mathbb{R} \), there is a unique \( \hat{\alpha}(t, \eta, x, p) \in U \) such that

\[
\hat{\alpha}(t, \eta, x, p) = \arg\min_{u \in U} h(t, \eta, x, u, p).
\]
From Berge’s maximum theorem (see [2, Theorem 17.31]) and part (b) of the above assumption it follows that \( \dot{\alpha} \) is a continuous function on \([0, T] \times \mathcal{P}([0, L]) \times [0, L] \times \mathbb{R} \). Also note that (3.10) implies that \( b, f, g, y, r \) are bounded functions, in particular,

\[
\sup_{(\eta, x, u) \in [0, T] \times \mathcal{P}([0, L]) \times [0, L] \times U} |b(t, \eta, x, u)| = c_B < \infty. \tag{3.10}
\]

The first part of the condition implies Hölder-1/2 continuity in \( t \) when \( \eta \) is replaced by \( \nu(t) \) for some \( \nu \in \mathcal{M}_0 \). For example, if \( \nu \in \mathcal{M}_c \), for \( t, t' \in [0, T] \)

\[
|y(t, \nu(t)) - y(t', \nu(t'))| \leq c_L (c + \sqrt{T})|t - t'|^{1/2}. \tag{3.11}
\]

Similar estimates hold for \( b, f, r \) and \( g \).

Assumption 3.1 will guarantee that the value function is the unique classical solution of the HJB equation (3.5) with BC (3.6) (see Proposition 3.1). The assumption will also be used in the fixed point argument (Section 3.3), in the asymptotic analysis of the \( n \)-player game (Section 4), and for numerical approximations (Section 5). Using Girsanov’s theorem it is easily checked that given a measurable function \( \gamma : [0, T] \times [0, L] \to U, \nu \in \mathcal{P}_{T,L} \), and \( (t, x) \in [0, T] \times [0, L] \) there is a unique weak solution \( Z \) to

\[
Z(s) = (X, Y, R)(s) = \Gamma \left( x + \int_0^s b(t + t', \nu(t + t'), X(t'), \gamma(t, X(t')))dt' + \sigma B(\cdot) \right)(s), \tag{3.12}
\]

\( s \in [0, T - t] \), namely there is a system \( \Xi = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P}, B) \) on which is given a \( \mathcal{F}_s \)-adapted continuous process \( Z = (X, Y, R) \) solving the above equation and if \( \Xi' = (\Omega', \mathcal{F}', \{\mathcal{F}'_t\}, \mathbb{P}', B') \) is another system on which is given a \( \mathcal{F}'_s \)-adapted continuous process \( Z' = (X', Y', R') \) solving the above equation with \( (Z, X, B) \) replaced by \( (Z', X', B') \) then \( \mathbb{P} \circ Z^{-1} = \mathbb{P}' \circ (Z')^{-1} \). Note also that with \( \alpha(s) = \gamma(s, X(s)), (\alpha, Z) \in \mathcal{A}(\Xi, t, x, \nu) \). We refer to the function \( \gamma \) as a feedback control and we call it an optimal feedback control for the stochastic control problem in (3.3) if

\[
V_\nu(t, x) = J_\nu(t, x, \alpha, Z).
\]

The following result which is a consequence of Theorem 13.24 of [53] says that the HJB equation (3.5) with BC (3.6) admits a unique classical solution which can be characterized as the value function \( V_\nu \). Moreover, there exists an optimal feedback control with certain regularity properties. We follow the notation from [53]. Let for \( \delta \in (0, 1], H_\delta \) be the collection of maps \( \psi : (0, T) \times (0, L) \to \mathbb{R} \) such that

\[
\sup_{0 < t' < t, 0 < x < x' < L} \frac{|\psi(t, x) - \psi(t, x')|}{|t - t'|^{\delta/2} + |x - x'|^{\delta}} < \infty.
\]

Note that such a function can be continuously extended to \([0, T] \times [0, L]\) and we denote the extension by the same symbol. Also, let \( H_{2+\frac{1}{\delta}} \) be the collection of continuous real functions \( \psi \) on \([0, T] \times [0, L]\) such that \( x \mapsto \psi(t, x) \) is twice continuously differentiable on \((0, L)\) for all \( t \in (0, T) \), \( t \mapsto \psi(t, x) \) is continuously differentiable on \((0, T)\) for all \( x \in (0, L) \), the functions

\[\psi, D_t \psi, D \psi, D^2 \psi,\]

are bounded on \((0, T) \times (0, L)\) and the functions \( D_t \psi, D^2 \psi \) are in \( H_{1/2} \).
Proposition 3.1 Fix $\nu \in \mathcal{M}_0$ and suppose that Assumption 3.1 holds. Then $V_\nu \in H_{2+\frac{1}{2}}$ and it is the unique solution of (3.5)–(3.6). Furthermore, with $\hat{\alpha}$ as introduced in Assumption 3.1, the map $(s,x') \mapsto \hat{\alpha}(s,\nu(s),x',D\nu(s,x'))$ is continuous and the feedback control $\hat{\gamma}(u,x') = \hat{\alpha}(u+t,\nu(u+t),x',D\nu(u+t,x'))$ is an optimal feedback control for (3.3) for every $t \in (0,T)$. Moreover, any optimal control $\alpha$ for (3.3) satisfies $\alpha(u,\omega) = \hat{\gamma}(u,X(u,\omega))$, $\lambda_T^\nu \otimes \mathbb{P}$ almost surely (a.s.), where $\lambda_T^\nu$ denotes the Lebesgue measure on $[0,T-t]$.

Proof: From [53, Theorem 13.24] and the paragraph following its statement it follows that (3.5)–(3.6) admits a solution in $H_{2+\frac{1}{2}}$. We remark that the key conditions needed to appeal to this theorem are that $\alpha \in M$ and suppose that Assumption 3.1 holds. Then $\hat{\alpha}$ is the unique solution of (3.5)–(3.6). Fix $(t,x) \in [0,T] \times [0,L]$, a system $\Xi$ and $(\alpha,Z) \in \mathcal{A}(\Xi,t,x,\nu)$. By an application of Itô’s lemma and using (3.5)–(3.6) we get

$$\mathbb{E}[\psi(T),X(T-t))] = \mathbb{E}[\psi(T),X(T-t))]
= \psi(t,x) + \mathbb{E} \left[ \int_0^{T-t} D\psi(s+t,0)dY(s) - \int_0^{T-t} D\psi(s+t,L)dR(s) \right] + \mathbb{E} \int_0^{T-t} Ds\psi(s+t,X(s)) + b(s+t,\nu(s+t),X(s),\nu(s))D\psi(s+t,X(s)) + \frac{1}{2} \sigma^2 D^2 \psi(s+t,X(s)) ds
\geq \psi(t,x) + \mathbb{E} \left[ - \int_0^{T-t} f(s+t,\nu(s+t),X(s),\nu(s)) ds - \int_0^{T-t} y(s+t,\nu(s+t))dY(s) - \int_0^{T-t} r(s+t,\nu(s+t))dR(s) \right].$$

Hence, with $\hat{\alpha}$ as introduced in Assumption 3.1, the map $(s,x') \mapsto \hat{\alpha}(s,\nu(s),x',D\nu(s,x'))$ is continuous and the feedback control $\hat{\gamma}(u,x') = \hat{\alpha}(u+t,\nu(u+t),x',D\nu(u+t,x'))$ is an optimal feedback control for (3.3) for every $t \in (0,T)$. Moreover, any optimal control $\alpha$ for (3.3) satisfies $\alpha(u,\omega) = \hat{\gamma}(u,X(u,\omega))$, $\lambda_T^\nu \otimes \mathbb{P}$ almost surely (a.s.), where $\lambda_T^\nu$ denotes the Lebesgue measure on $[0,T-t]$.

$$\mathbb{E}[\psi(T),X(T-t))] = \mathbb{E}[\psi(T),X(T-t))]
= \psi(t,x) + \mathbb{E} \left[ \int_0^{T-t} D\psi(s+t,0)dY(s) - \int_0^{T-t} D\psi(s+t,L)dR(s) \right] + \mathbb{E} \int_0^{T-t} Ds\psi(s+t,X(s)) + b(s+t,\nu(s+t),X(s),\nu(s))D\psi(s+t,X(s)) + \frac{1}{2} \sigma^2 D^2 \psi(s+t,X(s)) ds
\geq \psi(t,x) + \mathbb{E} \left[ - \int_0^{T-t} f(s+t,\nu(s+t),X(s),\nu(s)) ds - \int_0^{T-t} y(s+t,\nu(s+t))dY(s) - \int_0^{T-t} r(s+t,\nu(s+t))dR(s) \right].$$

Hence, with $Z = (X,Y,R),

$$J_\nu(t,x,\alpha,Z) = \mathbb{E} \left[ \int_0^{T-t} f(s+t,\nu(s+t),X(s),\nu(s)) ds + g(\nu(T),X(T-t)) + y(s+t,\nu(s+t))dY(s) + \int_0^{T-t} r(s+t,\nu(s+t))dR(s) \right]$$

$$\geq \psi(t,x).$$
Since $\Xi, \alpha, Z$ are arbitrary, we get that $V_\nu(t, x) \geq \psi(t, x)$ for all $(t, x) \in [0, T] \times [0, L]$. Let $\hat{\gamma}$ be as in the statement of the proposition with $V_\nu$ replaced by $\psi$. Let $Z = (X, Y, R)$ be a solution of (3.12) with $\gamma$ replaced by $\hat{\gamma}$ given on some system $\Xi$. Then a similar calculation using Itô’s formula but with $\alpha(u) = \hat{\gamma}(u, X(u))$ shows that $J_\nu(t, x, \alpha, Z) = \psi(t, x)$ for all $(t, x) \in [0, T] \times [0, L]$. This shows that $\psi = V_\nu$ and so $V_\nu$ is the unique $H_{2+\frac{1}{2}}$ solution of (3.5)–(3.6). Also, we have that $\hat{\gamma}$ as in the statement of the lemma is an optimal feedback control.

Since $V_\nu \in H_{2+\frac{1}{2}}$, the continuity of the map $(t, x) \mapsto \hat{\alpha}(t, \nu(t), x, DV_\nu(t, x))$ is immediate from the continuity of the map $\hat{\alpha}$ that was noted below Assumption 3.1.

We now show that $\hat{\gamma}$ is the unique optimal control for (3.3). Fix an optimal control $(\alpha, Z) \in \mathcal{A}(\Xi, t, x, \nu)$ given on some system $\Xi$. We claim that $\alpha(u, \omega) = \hat{\gamma}(u, X(u, \omega))$, $\lambda_T^x \times \mathbb{P}$-a.s. Indeed, suppose that there is a set with a positive $\lambda_T^x \times \mathbb{P}$-measure on which the equality fails. Then by Assumption 3.1(b) together with (3.5)–(3.6), it follows that (3.13) holds with a strict inequality, which in turn implies that (3.14) holds with a strict inequality. Recalling that $V_\nu = \psi$, we arrive at a contradiction. \(\square\)

### 3.3 Solving the MFG

We now turn to step (ii) in solving the fixed point problem. Although for existence of a fixed point Assumption 3.1 will suffice, in order to get uniqueness, we will need the following additional assumption. Similar assumption has been used to argue uniqueness of fixed points in previous works on MFG (see e.g. [51, Theorem 2.4], [20, Section 3], [31, equation (17)], [22, Assumption (U)]). Fix $\eta_0 \in \mathcal{P}([0, L])$.

**Assumption 3.2** For every $(t, \eta, x, u) \in [0, T] \times \mathcal{P}([0, L]) \times [0, L] \times U$, $b(t, \eta, x, u) = b(t, \eta_0, x, u)$, 
\[ f(t, \eta, x, u) = f_0(t, \eta, x) + f_1(t, x, u), \]  
\[ g(t, \eta) = g(t, \eta_0), \quad r(t, \eta) = r(t, \eta_0). \]  
Moreover, for every $t \in [0, T]$ and $\eta, \eta' \in \mathcal{P}([0, L])$, $f_0$ and $g$ satisfy the following monotonicity property
\[
\int_0^L [f_0(t, \eta, x) - f_0(t, \eta', x)]d(\eta - \eta')(x) \geq 0, \\
\int_0^L (g(\eta, x) - g(\eta', x))d(\eta - \eta')(x) \geq 0.
\]

Abusing notation, when Assumption 3.2 holds, we will write $b(t, x, u) = b(t, \eta_0, x, u)$, $y(t) = y(t, \eta_0)$, and $r(t) = r(t, \eta_0)$.

**Remark 3.1** Examples satisfying Assumption 3.2 are given in [20, page 8] and [17, page 6]. Another natural example that satisfies this assumption is a cost function that is linear in the mean-field term. That is,
\[ f_0(t, \eta, x) = a_1(t)(c_1 + \psi_1(x)) \int_0^L \psi_1(y)d\eta(y), \quad g(\eta, x) = (c_2 + \psi_2(x)) \int_0^L \psi_2(y)d\eta(y). \]
where \( a_1 : [0, T] \to \mathbb{R}_+ \), \( \psi_1, \psi_2 : [0, L] \to \mathbb{R} \), and \( c_1, c_2 \in \mathbb{R} \). From a modeling perspective, by choosing positive and nondecreasing \( \psi_i \)'s and a positive \( a_1 \), the system planner penalizes all servers collectively when the empirical measure has high \( \psi_i \)-moments and in addition it penalizes individual servers for long queues.

**Theorem 3.1** Under Assumption 3.7, there exists a solution of the MFG. If in addition Assumption 3.2 holds then there is a unique MFG solution.

The proof will appeal to Schauder’s fixed point theorem. Since Schauder’s original work (cf. [55]), there have been several versions of this result. We now quote the version ([50, Theorem 4.1.1]) that will be used here.

**Lemma 3.1** Let \( \mathcal{M} \) be a non-empty convex subset of a normed space \( \mathcal{B} \). Let \( \Phi \) be a continuous mapping of \( \mathcal{M} \) into a compact set \( \mathcal{K} \subset \mathcal{M} \). Then \( \Phi \) has a fixed point, namely there exists \( x \in \mathcal{K} \) such that \( \Phi(x) = x \).

**Proof of Theorem 3.1**: We will apply Lemma 3.1 to the space \( \mathcal{B} \) of finite signed measures on \( \mathcal{C}([0, T] : [0, L]) \) which is equipped with the Kantorovich-Rubinstein norm

\[
\|\nu\|_{KR} = \sup \left\{ \int_{\mathcal{C}([0, T] : [0, L])} F(\omega)d\nu(\omega) \mid F \in \text{Lip}_1(\mathcal{C}([0, T] : [0, L])), \nu \in \mathcal{B} \right\}.
\]

The distance driven by the norm, restricted to \( \mathcal{P}_{T,L} \), coincides with the Wasserstein’s distance of order 1 (see [59, Remark 6.5]), which due to compactness of \([0, L]\) generates the same topology on \( \mathcal{P}_{T,L} \) as that for weak convergence.

We now introduce a mapping \( \Phi \) on the non-empty and convex set \( \mathcal{M}_0 \) that satisfies the conditions stated in Lemma 3.1

**Definition of \( \Phi \)**. For \( \nu \in \mathcal{M}_0 \), let \( \alpha_\nu \) denote the optimal feedback control \( \hat{\gamma} \) for (3.3) (with \( t = 0 \)) given through Proposition 3.1. Let \( Z^\nu = (X^\nu, Y^\nu, R^\nu) \) denote the unique weak solution of (3.12) with \( \gamma \) replaced with \( \alpha_\nu \) given on some system \( \Xi = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P}, B) \). Define \( \Phi(\nu) = \mathbb{P} \circ (X^\nu)^{-1} \).

**Invariance of \( \mathcal{M}_0 \)**. For \( \nu \in \mathcal{M}_0 \) and \( 0 \leq s \leq t < T \)

\[
W_1(\Phi(\nu)(t), \Phi(\nu)(s)) \leq \mathbb{E}|X^\nu(t) - X^\nu(s)| \leq c_S \max\{c_B, \sigma\} \left( (t - s) + \mathbb{E} \sup_{s \leq u \leq t} |B(u) - B(s)| \right) \leq 4\tilde{C}(t - s)^{1/2},
\]

where \( \tilde{C} \equiv c_S \max\{c_B, \sigma\} \), and the last inequality uses Doob’s maximal inequality. This shows that \( \Phi(\nu) \in \mathcal{M}_0 \) for all \( \nu \in \mathcal{M}_0 \).

We now show that \( \Phi(\mathcal{M}_0) \) is contained in a compact set in \( \mathcal{P}_{T,L} \), i.e. \( \Phi(\mathcal{M}_0) \) is relatively compact.

**Relative compactness of \( \Phi(\mathcal{M}_0) \)**. For \( f \in \mathcal{C}([0, T] : \mathbb{R}) \) and \( \delta > 0 \), let

\[
\varphi_f(\delta) = \sup_{s, t \in [0, T] : |t - s| \leq \delta} |f(t) - f(s)|.
\]
Then for $\varepsilon, \delta, \rho > 0$ and $\nu \in \mathcal{M}_0$, similar to the estimate in (3.17),

$$\Phi(\nu)(X^\nu : \varpi_X(\delta) \geq \rho) \leq \frac{1}{\rho} \mathbb{E} \sup_{0 \leq s \leq t + \delta \leq T} |X^\nu(t) - X^\nu(s)|$$

$$\leq \frac{C}{\rho} \left( \delta + \mathbb{E} \sup_{0 \leq s \leq t + \delta \leq T} \sup_{s \leq u \leq t} |B(u) - B(s)| \right)$$

$$\leq \frac{C}{\rho} (\delta + \mathbb{E} \varpi_B(\delta)).$$

Since $\mathbb{E} \varpi_B(\delta) \to 0$ as $\delta \to 0$, we have from [12] Theorem 7.3 that $\{\Phi(\nu) : \nu \in \mathcal{M}_0\}$ is relatively compact in $\mathcal{P}_{T,L}$.

**Continuity of $\Phi$.** We now argue that $\Phi$ is a continuous map on $\mathcal{M}_0$. Consider a system $\Xi = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P}, B)$ and let $Z^0 = (X^0, Y^0, R^0)$ be given as

$$Z^0(s) = (X^0(s), Y^0(s), R^0(s)) = \Gamma(x + \sigma B(\cdot))(s), \ 0 \leq s \leq T. \quad (3.18)$$

Define for $\nu' \in \mathcal{M}_0$

$$\alpha_{\nu'}^0(t, X^0) = \hat{\alpha}(t, \nu'(t), X^0(t), \nu'(t), X^0(t)), \ t \in [0, T]. \quad (3.19)$$

Let $\nu^n, \nu \in \mathcal{M}_0$ be such that $\nu^n \to \nu$. Since $\sup_{0 \leq t \leq T} W_1(\nu^n(t), \nu(t)) \leq W_1(\nu^n, \nu)$, the above convergence implies that

$$\lim_{n \to \infty} \sup_{0 \leq t \leq T} W_1(\nu^n(t), \nu(t)) = 0. \quad (3.20)$$

We now show that $\alpha^n_{\nu_n} \to \alpha^0_{\nu}$ in $\lambda_T \times \mathbb{P}$-measure. Recall that $\hat{\alpha}$ is a continuous map. Hence, in view of (3.20), for the desired convergence, it is sufficient to show that

$$\lim_{n \to \infty} \mathbb{E} \left[ \int_0^T |DV_{\nu^n}(t, X^0(t)) - DV_{\nu}(t, X^0(t))| dt \right] = 0. \quad (3.21)$$

The proof is a modification of the proof of Theorem 2.1 in [37]. Applying Itô's lemma to $V_{\nu'}(t, X^0(t))$ for fixed $\nu' \in \mathcal{M}_0$ and using from Proposition 3.1 the fact that $V_{\nu'}$ solves (3.5)–(3.6), we have for every $t \in [0, T]$,

$$V_{\nu'}(t, X^0(t)) - V_{\nu'}(T, X^0(T)) = \int_t^T H(s, \nu'(s), X^0(s), DV_{\nu'}(s, X^0(s))) ds - \sigma \int_t^T DV_{\nu'}(s, X^0(s)) dB(s)$$

$$+ \int_t^T y(s, \nu'(s)) dB(s) + \int_t^T r(s, \nu'(s)) dR^0(s). \quad (3.22)$$

Let,

$$\Delta V^n(t) = V_{\nu^n}(t, X^0(t)) - V_{\nu}(t, X^0(t)), \ \Delta DV^n(t) = DV_{\nu^n}(t, X^0(t)) - DV_{\nu}(t, X^0(t))$$

$$\Delta g^n(T) = \Delta V^n(T) = g(\nu^n(T), X^0(T)) - g(\nu(T), X^0(T)),$$

$$\Delta y^n(t) = y(t, \nu^n(t)) - y(t, \nu(t)), \ \Delta r^n(t) = r(t, \nu^n(t)) - r(t, \nu(t)),$$

$$\Delta \psi^n(t) = \sup_{\alpha \in U} |\psi(t, \nu^n(t), X^0(t), \alpha) - \psi(t, \nu(t), X^0(t), \alpha)|, \ \psi \in \{f, b\}. \quad (3.23)$$
Then, (3.22) and (3.23) imply,

\[
\Delta V^n(t) + \sigma \int_t^T \Delta V^n(s) dB(s) \\
= \Delta g^n(T) + \int_t^T \Delta y^n(s) dY^0(s) + \int_t^T \Delta r^n(s) dR^0(s) \\
+ \int_t^T [H(s, \nu^n(s), X^0(s), D\nu(s, X^0(s))) - H(s, \nu^n(s), X^0(s), D\nu(s, X^0(s))))] ds \\
+ \int_t^T [H(s, \nu^n(s), X^0(s), D\nu(s, X^0(s))) - H(s, \nu(s), X^0(s), D\nu(s, X^0(s))))] ds.
\]

Squaring both sides and then taking expectations gives

\[
\mathbb{E}[(\Delta V^n(t))^2] + \sigma^2 \mathbb{E} \left[ \int_t^T (\Delta V^n(s))^2 ds \right] \tag{3.24}
\leq 2(T - t) \mathbb{E} \left[ \int_t^T [H(s, \nu^n(s), X^0(s), D\nu(s, X^0(s))) - H(s, \nu^n(s), X^0(s), D\nu(s, X^0(s))))]^2 ds \right] \\
+ 2C^n(t) \\
\leq 2c_B^2 (T - t) \mathbb{E} \left[ \int_t^T (\Delta V^n(s))^2 ds \right] + 2C^n(t),
\]

where

\[
C^n(t) = \mathbb{E} \left[ \Delta g^n(T) + \int_t^T \Delta y^n(s) dY^0(s) + \int_t^T \Delta r^n(s) dR^0(s) \\
+ \int_t^T [H(s, \nu^n(s), X^0(s), D\nu(s, X^0(s))) - H(s, \nu^n(s), X^0(s), D\nu(s, X^0(s))))] ds \right]^2.
\]

Letting \( \delta = \sigma^2 / (4c_B^2) \), we get from (3.24) that for every \( t \in [T - \delta, T] \),

\[
\mathbb{E}[(\Delta V^n(t))^2] + \frac{\sigma^2}{2} \mathbb{E} \left[ \int_t^T (\Delta V^n(s))^2 ds \right] \leq 4C^n(t). \tag{3.25}
\]

We now show that \( \limsup_{n \to \infty} \sup_{0 \leq t \leq T} C^n(t) = 0 \). Note that there exists \( C_1 \in (0, \infty) \) such that the following inequality holds for all \( n \in \mathbb{N} \).

\[
\frac{1}{C_1} \sup_{0 \leq t \leq T} C^n(t) \leq \mathbb{E} \left[ (\Delta g^n(T))^2 \right] + \mathbb{E} \left[ (Y^0(T))^2 \right] \sup_{0 \leq t \leq T} (\Delta y^n(t))^2 + \mathbb{E} \left[ (R^0(T))^2 \right] \sup_{0 \leq t \leq T} (\Delta r^n(t))^2 \\
+ \mathbb{E} \left[ \int_0^T (\Delta f^n(s))^2 ds \right] + \mathbb{E} \left[ \int_0^T (\Delta b^n(s))^2 ds \right] \int_0^T (D\nu(s, X^0(s)))^2 ds. \tag{3.26}
\]

Using the properties of the Skorohod map (Lemma 2.1), it follows that

\[
\mathbb{E} \left[ (Y^0(T))^2 + (R^0(T))^2 \right] < \infty.
\]
Using the above property we will now argue that \( \Phi(\nu^n) \to \Phi(\nu) \) as \( n \to \infty \), completing the proof of continuity of \( \Phi \). Let for \( \nu' \in \mathcal{M}_0 \),

\[
\gamma_{\nu'}(t, x) = \hat{\alpha}(t, \nu'(t), x, DV_{\nu'}(t, x)), \quad \hat{b}_{\nu'}(t, x) = b(t, \nu'(t), x, \gamma_{\nu'}(t, x)), \quad (t, x) \in [0, T] \times [0, L]
\]

and let \( \mathbb{P}^{\nu'} \) be a probability measure on \( (\Omega, \mathcal{F}) \) defined as

\[
d\mathbb{P}^{\nu'} = \exp \left[ \frac{1}{\sigma} \int_0^T \hat{b}_{\nu'}(t, X^0(t)) dB(t) - \frac{1}{2\sigma^2} \int_0^T \hat{b}_{\nu'}^2(t, X^0(t)) dt \right] d\mathbb{P}. \tag{3.29}
\]

By Girsanov’s theorem \( \mathbb{P}^{\nu'} \circ (X^0)^{-1} = \mathbb{P}(\nu') \). Thus to show \( \Phi(\nu^n) \to \Phi(\nu) \) it suffices to argue that \( \mathbb{P}^{\nu^n} \to \mathbb{P}^{\nu} \). We will in fact show that \( R(\mathbb{P}^{\nu'}||\mathbb{P}^{\nu}) \to 0 \) as \( n \to \infty \), where

\[
R(\mathbb{P}^{\nu'}||\mathbb{P}^{\nu}) = \mathbb{E}^{\nu} \left( \log \frac{d\mathbb{P}^{\nu'}}{d\mathbb{P}^{\nu}} \right) = \mathbb{E} \left( \frac{d\mathbb{P}^{\nu'}}{d\mathbb{P}} \log \frac{d\mathbb{P}^{\nu'}}{d\mathbb{P}^{\nu}} \right)
\]

is the relative entropy of \( \mathbb{P}^{\nu} \) with respect to \( \mathbb{P}^{\nu^n} \), which due to Pinsker’s inequality (see [58, Page 132]) gives the convergence of \( \mathbb{P}^{\nu^n} \) to \( \mathbb{P}^{\nu} \).

Let

\[
\Delta \hat{b}_n(t, x) = \hat{b}_{\nu}(t, x) - \hat{b}_{\nu^n}(t, x), \quad \Delta \hat{b}_n^2(t, x) = \hat{b}_{\nu}^2(t, x) - \hat{b}_{\nu^n}^2(t, x), \quad (t, x) \in [0, T] \times [0, L].
\]

With this notation,

\[
\log \frac{d\mathbb{P}^{\nu'}}{d\mathbb{P}^{\nu^n}} = \frac{1}{\sigma} \int_0^T \Delta \hat{b}_n(t, X^0(t)) dB(t) - \frac{1}{2\sigma^2} \int_0^T \Delta \hat{b}_n^2(t, X^0(t)) dt.
\]

Also, noting that since \( b \) is bounded \( \mathbb{E}(\frac{d\mathbb{P}^{\nu'}}{d\mathbb{P}})^2 \preceq \kappa < \infty \), we have from Cauchy-Schwarz inequality

\[
R(\mathbb{P}^{\nu'}||\mathbb{P}^{\nu}) \leq \sqrt{\mathbb{E} \left[ \frac{1}{\sigma} \int_0^T \Delta \hat{b}_n(t, X^0(t)) dB(t) - \frac{1}{2\sigma^2} \int_0^T \Delta \hat{b}_n^2(t, X^0(t)) dt \right]^2} \tag{3.30}
\]

Next note that

\[
\mathbb{E} \left[ \frac{1}{\sigma} \int_0^T \Delta \hat{b}_n(t, X^0(t)) dB(t) \right]^2 = \frac{1}{\sigma^2} \int_0^T \mathbb{E} [\Delta \hat{b}_n(t, X^0(t))]^2 dt = \frac{1}{\sigma^2} \int_0^T \mathbb{E} [b(t, \nu(t), X^0(t), \alpha_0^\nu(t, X^0(t))) - b(t, \nu^n(t), X^0(t), \alpha_0^\nu(t, X^0(t)))]^2 dt.
\]
The last term converges to 0 from the boundedness and continuity of \( b \), (3.20) and (3.28). Similarly
\[
\mathbb{E} \left[ \frac{1}{2\sigma^2} \int_0^T \Delta h^2(t, X^0(t)) dt \right]^2 \to 0
\]
as \( n \to \infty \). Using the above two observations in (3.30) we have \( R(\mathbb{P}^\nu \| \mathbb{P}^\nu) \to 0 \) and thus, as argued earlier, the proof of continuity of \( \Phi \) is complete. Thus we have shown that \( \Phi \) is a continuous map on \( \mathcal{M}_0 \), which is a non-empty convex subset of the normed space \( \mathcal{B} \), into a compact set \( \mathcal{K} \subset \mathcal{M}_0 \). Thus by the fixed point theorem in Lemma 3.1 \( \Phi \) has a fixed point.

The first results on unique solvability of a MFG go back to [51]. Since then uniqueness has been argued in various settings (see e.g. [20, 31, 22]). The proof given below uses arguments similar to those in [22, Section 7.3], however for the sake of completeness we give the details. Consider as before a system \( \Xi = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P}, B) \) and let \( Z^0 = (X^0, Y^0, R^0) \) be given through (3.18). Let \( \nu_1, \nu_2 \in \mathcal{M}_0 \). For \( i = 1, 2 \), let \( \alpha^0_{\nu_i} \) and \( \mathbb{P}^{\nu_i} \) be given by (3.19) and (3.29), respectively, with \( \nu_1 \) replacing \( \nu' \). Applying Itô’s lemma to \( V_{\nu_1}(t, X^0(t)) \) and recalling that \( V_{\nu_1} \) solves (3.5)–(3.6) with \( \nu \) replaced with \( \nu_1 \), we get
\[
V_{\nu_1}(T, X^0(T)) - V_{\nu_1}(0, x) = -\int_0^T H(t, \nu_1(t), X^0(t), DV_{\nu_1}(t, X^0(t))) dt + \sigma \int_0^T DV_{\nu_1}(t, X^0(t)) dB(t) + \zeta_T
\]
where recalling the form of \( y \) and \( r \) from Assumption 3.2
\[
\zeta_T = \int_{[0,T]} DV_{\nu_1}(t, 0) dY^0(t) + \int_{[0,T]} DV_{\nu_1}(t, 0) dR^0(t)
\]
\[
= \int_{[0,T]} y(t) dY^0(t) + \int_{[0,T]} r(t) dR^0(t).
\]
Observing that for \( t \in [0, T] \)
\[
h(t, \nu_1(t), X^0(t), \alpha^0_{\nu_1}(t), DV_{\nu_1}(t, X^0(t))) = h(t, \nu_1(t), X^0(t), \alpha^0_{\nu_1}(t), DV_{\nu_2}(t, X^0(t))) + \mathcal{R}_t,
\]
where
\[
\mathcal{R}_t = [DV_{\nu_1}(t, X^0(t)) - DV_{\nu_2}(t, X^0(t))] \tilde{h}_{\nu_1}(t, X^0(t)),
\]
we have
\[
V_{\nu_1}(0, x) = g(\nu_1(T), X^0(T)) + \int_0^T h(t, \nu_1(t), X^0(t), \alpha^0_{\nu_1}(t), DV_{\nu_2}(t, X^0(t))) dt
\]
\[
+ \int_0^T \mathcal{R}_t dt - \sigma \int_0^T DV_{\nu_1}(t, X^0(t)) dB(t) - \zeta_T. \tag{3.31}
\]
Similarly, applying Itô’s lemma to $V_{\nu_2}(t, X^0(t))$

\[
V_{\nu_2}(0, x) = g(\nu_2(T), X^0(T)) + \int_0^T H(t, \nu_2(t), X^0(t), DV_{\nu_2}(t, X^0(t)))dt - \sigma \int_0^T DV_{\nu_2}(t, X^0(t))d\nu(t) - \zeta_T.
\]

Subtracting (3.32) from (3.31)

\[
V_{\nu_1}(0, x) - V_{\nu_2}(0, x)
= g(\nu_1(T), X^0(T)) - g(\nu_2(T), X^0(T)) - \sigma \int_0^T [DV_{\nu_1}(t, X^0(t)) - DV_{\nu_2}(t, X^0(t))]d\nu(t)
+ \int_0^T [h(t, \nu_1(t), X^0(t), \alpha_{\nu_1}(t), DV_{\nu_2}(t, X^0(t))) - H(t, \nu_2(t), X^0(t), DV_{\nu_2}(t, X^0(t))))] dt
+ g(\nu_1(T), X^0(T)) - g(\nu_2(T), X^0(T))
\]

where for $i = 1, 2$, $B^\nu_i(t) \equiv B(t) - \frac{1}{\sigma} \int_0^t \dot{b}_{\nu_i}(s, X^0(s))ds, t \in [0, T]$. Since under $\mathbb{P}^{\nu_i}$, $B^\nu_i$ is a standard Brownian motion, taking expectation under the measure $\mathbb{P}^{\nu_i}$

\[
V_{\nu_1}(0, x) - V_{\nu_2}(0, x)
= \mathbb{E}_{\nu_1}\left[\int_0^T [h(t, \nu_1(t), X^0(t), \alpha_{\nu_1}(t), DV_{\nu_2}(t, X^0(t))) - H(t, \nu_2(t), X^0(t), DV_{\nu_2}(t, X^0(t))))] dt
+ g(\nu_1(T), X^0(T)) - g(\nu_2(T), X^0(T))\right].
\]

A similar calculation shows

\[
V_{\nu_1}(0, x) - V_{\nu_2}(0, x)
= \mathbb{E}_{\nu_2}\left[\int_0^T [H(t, \nu_1(t), X^0(t), DV_{\nu_1}(t, X^0(t))) - h(t, \nu_2(t), X^0(t), \alpha_{\nu_2}(t), DV_{\nu_1}(t, X^0(t))))] dt
+ g(\nu_1(T), X^0(T)) - g(\nu_2(T), X^0(T))\right].
\]

By the definition of the $H$ and the form of $f$ in Assumption 3.2 we get,

\[
\begin{align*}
\begin{array}{ll}
h(s, \nu_1(s), X^0(s), \alpha_{\nu_1}(s), DV_{\nu_2}(s, X^0(s))) \\
\geq H(s, \nu_1(s), X^0(s), DV_{\nu_2}(s, X^0(s))) \\
= f_0(s, \nu_1(s), X^0(s)) + f_1(s, X^0(s), \alpha_{\nu_2}(s)) + DV_{\nu_2}(s, X^0(s))\dot{b}_{\nu_2}(s)
\end{array}
\end{align*}
\]

where the last equality uses the observation that since $\dot{\alpha}(s, \eta, x, p)$ does not depend on $\eta$,

\[
\dot{\alpha}(s, \nu_1(s), X^0(s), DV_{\nu_2}(s, X^0(s))) = \alpha_{\nu_2}(s).
\]

Therefore for all $s \in [0, T],$

\[
\begin{align*}
\begin{array}{ll}
h(s, \nu_1(s), X^0(s), \alpha_{\nu_1}(s), DV_{\nu_2}(s, X^0(s))) - H(s, \nu_2(s), X^0(s), DV_{\nu_2}(s, X^0(s))) \\
\geq f_0(s, \nu_1(s), X^0(s)) - f_0(s, \nu_2(s), X^0(s))
\end{array}
\end{align*}
\]

Similarly for all $s \in [0, T],$

\[
\begin{align*}
\begin{array}{ll}
H(s, \nu_1(s), X^0(s), DV_{\nu_1}(s, X^0(s))) - h(s, \nu_2(s), X^0(s), \alpha_{\nu_2}(s), DV_{\nu_1}(s, X^0(s))) \\
\leq f_0(s, \nu_1(s), X^0(s)) - f_0(s, \nu_2(s), X^0(s))
\end{array}
\end{align*}
\]

(3.35)
Applying the last two inequalities to (3.33) and (3.34), we get
\[ 0 \leq [\mathbb{E}^\nu - \mathbb{E}^\mu] \left[ g(\nu_1(T), X^0(T)) - g(\nu_2(T), X^0(T)) \right] + \int_0^T \left[ f_0(s, \nu_1(s), X^0(s)) - f_0(s, \nu_2(s), X^0(s)) \right] ds. \] (3.37)

Until now \( \nu_1 \) and \( \nu_2 \) were arbitrary measures in \( \mathcal{M}_0 \). Suppose now that \( \nu_i, i = 1, 2 \), are fixed points of \( \Phi \). Then, for \( i = 1, 2 \), \( \mathbb{P}^{\nu_i} \circ (X^0)^{-1} = \nu_i \) and so for all \( s \in [0, T] \), \( \mathbb{P}^{\nu_i} \circ (X^0(s))^{-1} = \nu_i(s) \).

In this case, using the inequalities in Assumption 3.2 we get that the inequality (3.37) can be replaced with equality. We claim that \( \alpha_{\nu_i}^0 = \alpha_{\nu_i}^0, \lambda_T^0 \times \mathbb{P}\text{-a.s.} \) Indeed, suppose that there is a set with positive \( \lambda_T^0 \times \mathbb{P}\)-measure on which \( \alpha_{\nu_i}^0 \neq \alpha_{\nu_i}^0 \). Then on this set (3.33) and (3.34) will hold with strict inequalities by Assumption 3.1 (b). Since the measures \( \mathbb{P}, \mathbb{P}^{\nu_1}, \) and \( \mathbb{P}^{\nu_2} \) are equivalent, this will say that (3.37) holds with a strict inequality as well, which contradicts the equality that was established above. This proves the claim. Since \( b(t, \eta, x, u) \) does not depend on \( \eta \) we conclude from the equality of \( \alpha_{\nu_1}^0 \) and \( \alpha_{\nu_2}^0 \) that \( \hat{\nu}_1(t, X^0(t, \omega)) = \hat{\nu}_2(t, X^0(t, \omega)) \), \( \lambda_T^0 \times \mathbb{P}\text{-a.s.} \) and thus by (3.29), \( \mathbb{P}^{\nu_1} = \mathbb{P}^{\nu_2} \). Combining this with the fact that \( \nu_i \) are fixed points of \( \Phi \), we now have \( \nu_1 = \mathbb{P}^{\nu_1} \circ (X^0)^{-1} = \mathbb{P}^{\nu_2} \circ (X^0)^{-1} = \nu_2 \).

**4 Asymptotic Nash equilibrium**

The main result of this section is Theorem 4.1. The main idea in the proof is to use a solution \( \bar{\nu} \) to the MFG (which from Theorem 3.1 exists under Assumption 3.1) and the associated feedback control given by Proposition 3.1 in order to construct an admissible control \( \bar{\alpha}^n = \{\bar{\alpha}^{n,i}\}_{i=1}^n \) for the n-player game. Specifically, the control will be given in a feedback form through the following relation
\[ \bar{\alpha}^{n,i}(t) \doteq \hat{\alpha}(t, \bar{\nu}(t), \bar{Q}^{n,i}(t), DV_{\bar{\nu}}(t, \bar{Q}^{n,i}(t)) \right), \] (4.1)

where \( \bar{Q}^{n,i} \) is the corresponding scaled queue length under the feedback control. Note that the only information each of the players uses is its own state. Therefore, the problem is decentralized in the sense that players do not need to observe each others’ states. Our main condition, in addition to Assumption 3.1 for \( \{\bar{\alpha}^n\} \) to be an asymptotic Nash equilibrium is the following. It in particular says that the drift function does not depend on the mean-field term. Fix \( \eta_0 \in \mathcal{P}([0, L]) \).

**Assumption 4.1**

(a) For every \( t \in [0, T], \eta \in \mathcal{P}([0, L]), \ x \in [0, L] \), and \( \alpha \in U \), one has,
\[ b(t, \eta, x, \alpha) = b(t, \eta_0, x, \alpha); \]

(b) There exists \( x \in [0, L] \) such that for every \( i \in \mathbb{N} \),
\[ \lim_{n \to \infty} \bar{Q}^{n,i}(0) = x. \] (4.2)
As before, with an abuse of notation, we will write $b(t, \eta, x, \alpha)$ as $b(t, x, \alpha)$ when Assumption 4.1 holds. As discussed in Remark 3.1 part (a) of the assumption means that the empirical measure affects the drift only through the control, which in turn is affected by the empirical measure through the running cost.

Recall the probability space $(\Omega', F', P')$ from Section 2.1. Let for $n \in \mathbb{N}$, $t \in [0, T]$ and $i = 1, \ldots, n$, $\beta^n(t) : \Omega' \to U$ be such that

$$\tilde{\alpha}^{-i}_n = \{\alpha^{n,1}, \ldots, \alpha^{n,i-1}, \beta^n, \alpha^{n,i+1}, \ldots, \alpha^{n,n}\}$$

is an admissible control (i.e. $\tilde{\alpha}^{-i}_n \in \mathcal{U}^n$). The following is the main result of the section.

**Theorem 4.1** Suppose Assumptions 3.1 and 4.1 hold. Let $\tilde{\nu}_n$, $\tilde{\alpha}_n$ and $\tilde{\alpha}^{-i}_n$, $i = 1, \ldots, n$, $n \in \mathbb{N}$ be as introduced above. Then

$$\limsup_{n \to \infty} J^{n,1}_n(\tilde{Q}_n(0); \tilde{\alpha}_n) = V_{\tilde{\nu}}(0, x) \leq \liminf_{n \to \infty} J^{n,1}_n(\tilde{Q}_n(0); \tilde{\alpha}^{-i}_n). \quad (4.3)$$

Theorem 4.1 in particular says that for every $\varepsilon > 0$, there is $n \in \mathbb{N}$ sufficiently large such that $\tilde{\alpha}^n$ forms an $\varepsilon$-Nash equilibrium in the $n$-player game.

The proof is given in the next three sections. First in Section 4.1 (Proposition 4.1) we will prove the convergence of empirical measures of the scaled queue length processes under controls $\tilde{\alpha}_n$ and $\tilde{\alpha}^{-i}_n$ to $\tilde{\nu}$. Next, in Section 4.2 (Proposition 4.2) we will prove the first equality in (4.3) and finally Proposition 4.3 in Section 4.3 will prove the inequality in (4.3).

### 4.1 Convergence of empirical measures

Let for $i \in \mathbb{N}$, $\tilde{\alpha}^{-i}_n \in \mathcal{U}^n$ be as defined below Assumption 4.1. Let

$$\tilde{\nu}^{-i}_n = \frac{1}{n} \sum_{j=1}^{n} \delta_{\tilde{Q}^{-i}_n(j)},$$

where $\tilde{Q}^{-i}_n(j)$ is the controlled queue length process defined by (2.5) with $\alpha^n$ replaced with $\tilde{\alpha}^{-i}_n$. The following result gives the convergence of $\tilde{\nu}^{-i}_n$ to $\tilde{\nu}$.

**Proposition 4.1** Suppose Assumptions 3.1 and 4.1 hold. Then for every $i \in \mathbb{N}$, $\tilde{\nu}^{-i}_n$ converges in probability, in $\mathcal{P}(\mathcal{D}([0, T] : [0, L]))$, to $\tilde{\nu}$ as $n \to \infty$.

**Proof:** Without loss of generality we assume that $i = 1$. Recall that for $j \in \mathbb{N}$, $\tilde{Q}^{n,j}$ is defined by (2.5) with $\alpha^n$ replaced with $\tilde{\alpha}^{-1}_n$. Define

$$\tilde{\zeta}^{n,i}(t) = \tilde{Q}^{n,i}(0) + \tilde{A}^{n,i}(t) - \tilde{D}^{n,i}(t) + \int_0^t \tilde{b}^{n,i}(s) ds, \quad t \in [0, T],$$

where $\tilde{A}^{n,i}, \tilde{D}^{n,i}$ are as in (2.3). Define for $i = 1, \ldots, n$,

$$\tilde{G}^{n,i} = (\zeta^{n,i}, \tilde{Q}^{n,i}, \tilde{Y}^{n,i}, \tilde{R}^{n,i})$$

and let

$$\Xi^n = \frac{1}{n-1} \sum_{i=2}^{n} \delta_{\tilde{G}^{n,i}}.$$
Note that since by assumption \( \{ \tilde{Q}^{n,i}(0) \}_{i=1}^{n} \) are exchangeable and the controls \( \tilde{\alpha}^{n,i} \) are given in terms of the same feedback function \( \tilde{\alpha} \) for each \( i = 2, \ldots, n \), the processes \( \{ \tilde{G}^{n,i} \}_{i=2}^{n} \) are exchangeable. Defining \( \tilde{G}^{n,i} \) to be the zero process for \( i > n \) we can regard, \( \tilde{G}^{n} = \{ \tilde{G}^{n,i} \}_{i=2}^{n} \) as a random variable with values in \( D([0,T] : (\mathbb{R}^{4})^{{\otimes}\infty}) \). We now argue the tightness of the sequence \( \{ \tilde{G}^{n} \} \). It suffices to show for each \( i \), the tightness of \( \{ \tilde{G}^{n,i} \}_{n\in\mathbb{N}} \) in \( D([0,T] : \mathbb{R}^{4}) \). Since
\[
\sup_{n,i,\omega} \sup_{t \in [0,T]} \frac{\lambda^{n,i}(t,\omega) + \mu^{n,i}(t,\omega)}{e_{n}} \equiv C_{0} < \infty, \tag{4.4}
\]
the following two conditions are satisfied with \( X^{n} \) equal to \( \langle \tilde{A}^{n,i}, \tilde{A}^{n,j} \rangle \) and \( \langle \tilde{D}^{n,i}, \tilde{D}^{n,j} \rangle \) for all \( i, j \).

[A] For each \( \varepsilon > 0, \eta > 0 \) there exists a \( \delta > 0 \) and \( n_{0} \in \mathbb{N} \) with the property that for every family of stopping times \( \{ \tau_{n} \}_{n\in\mathbb{N}} \) (\( \tau_{n} \) being an \( \mathcal{F}_{\tau} \)-stopping time on \( (\Omega', \mathcal{F}', \mathbb{P}') \)) with \( \tau_{n} \leq T - \delta \),
\[
\sup_{n \geq n_{0}} \sup_{\theta \leq \delta} \mathbb{P}\{ \| X^{n}(\tau_{n}) - X^{n}(\tau_{n} + \theta) \| \geq \eta \} \leq \varepsilon.
\]

[T\textsubscript{1}] For every \( t \) in some dense subset of \( [0,T] \), \( \{ X^{n}(t) \}_{n\in\mathbb{N}} \) is a tight sequence of \( \mathbb{R} \) valued random variables.

Then by Rebolledo’s theorem (see \[41\] Theorem 2.3.2) \( \{ \tilde{A}^{n,i} \}_{n\geq1} \) and \( \{ \tilde{D}^{n,i} \}_{n\geq1} \) are tight in \( D([0,T] : \mathbb{R}) \) for each \( i \). Also since the jumps of \( A^{n,i} \) and \( D^{n,i} \) are of size \( 1/\sqrt{e_{n}} \), these processes are \( \mathcal{C} \)-tight (namely all weak limit points are continuous a.s.). From boundedness of \( b \) we see that \( \{ \int_{0}^{s} \tilde{b}^{n,i}(s)ds \}_{n\in\mathbb{N}} \) is tight in \( \mathcal{C}([0,T] : \mathbb{R}) \). Combining this with Assumption \[41\] (b) we see that \( \{ \tilde{\zeta}^{n,i} \}_{n\in\mathbb{N}} \) is \( \mathcal{C} \)-tight in \( D([0,T] : \mathbb{R}) \). Using now the continuity of the Skorohod map (Lemma 2.4) we have the desired tightness of \( \{ \tilde{G}^{n,i} \}_{n\in\mathbb{N}} \).

Suppose now that, along some subsequence, \( \tilde{G}^{n} \) converges to \( \tilde{G} = (\zeta, \tilde{\zeta}^{i}, \tilde{Y}^{i}, \tilde{R}^{i}) \), in distribution, in \( D([0,T] : (\mathbb{R}^{4})^{{\otimes}\infty}) \). Then \( \tilde{G} \in \mathcal{C}([0,T] : (\mathbb{R}^{4})^{{\otimes}\infty}) \) a.s. and from \[42\] Lemma 4.2 and the exchangeability of \( \{ \tilde{G}^{n,i} \}_{i=2}^{n} \) it follows that \( \{ \tilde{G}^{i} \}_{i=2}^{n} \) is exchangeable and (along the subsequence),
\[
(\tilde{G}^{n}, \Xi^{n}) \Rightarrow (\tilde{G}, \Xi) \tag{4.5}
\]
in \( D([0,T] : (\mathbb{R}^{4})^{{\otimes}\infty}) \times \mathcal{P}(D([0,T] : \mathbb{R}^{4})) \) where \( \Xi \doteq \lim_{m \to \infty} \frac{1}{m-1} \sum_{i=2}^{m} \delta_{\tilde{G}^{i}} \).

We will now characterize the distribution of \( \{ \tilde{Q}^{i} \} \). From tightness of \( \{ \tilde{Y}^{n,i} \}_{n\in\mathbb{N}} \) and \( \{ \tilde{R}^{n,i} \}_{n\in\mathbb{N}} \) argued above and \[2.6\] it follows that
\[
\frac{1}{e_{n}} \int_{0}^{t} 1_{\{ \tilde{Q}^{n,i}(s) = 0 \}} \mu^{n,i}(s)ds \to 0, \quad \frac{1}{e_{n}} \int_{0}^{t} 1_{\{ \tilde{Q}^{n,i}(s) = L \}} \lambda^{n,i}(s)ds \to 0,
\]
uniformly on \( [0,T] \), in probability. Also, from \[2.7\] it follows that
\[
\sup_{0 \leq s \leq T} \left[ \frac{\lambda^{n,i}(s)}{e_{n}} - \lambda \right] + \left[ \frac{\mu^{n,i}(s)}{e_{n}} - \mu \right] \to 0 \text{ a.s.}
\]
as \( n \to \infty \). Thus from \[2.8\] (and the relation \( \lambda = \mu \)), for all \( i, j \),
\[
(\tilde{A}^{n,i}, \tilde{A}^{n,j})(t) \to \delta_{ij} \lambda, \quad (\tilde{D}^{n,i}, \tilde{D}^{n,j})(t) \to \delta_{ij} \lambda, \quad (\tilde{A}^{n,i}, \tilde{D}^{n,j})(t) \to 0
\]
in probability, uniformly on \([0, T]\), as \(n \to \infty\). By standard martingale techniques it now follows that

\[
\{\tilde{A}^{n,i} - \tilde{D}^{n,i}\}_{i \geq 1} \Rightarrow \{\sigma B^i\}_{i \geq 1},
\]

in \(D([0, T] : \mathbb{R}^\infty)\), where \(\{B^i\}\) are mutually independent standard Brownian motions. Also, since for \(i \geq 2\)

\[
(\tilde{Q}^{n,i}, \tilde{Y}^{n,i}, \tilde{R}^{n,i})(t) = \Gamma \left( \tilde{Q}^{n,i}(0) + \int_0^t \tilde{b}^{n,i}(s) ds + \tilde{A}^{n,i}(-) - \tilde{D}^{n,i}(-) + o(1) \right)(t), \ t \in [0, T],
\]

where

\[
\tilde{b}^{n,i}(t) = b(t, \tilde{Q}^{n,i}(t), \hat{\alpha}(t, \tilde{\nu}(t), \tilde{Q}^{n,i}(t), DV_\nu(t, \tilde{Q}^{n,i}(t))),
\]

we have from the continuity of \(b\) (Assumption 3.1), \(\hat{\alpha}\), and \(DV_\nu\), for \(i \geq 2\),

\[
(\tilde{Q}^i, \tilde{Y}^i, \tilde{R}^i)(t) = \Gamma \left( x + \int_0^t b(t, \tilde{Q}^i(t), \hat{\alpha}(t, \tilde{\nu}(t), \tilde{Q}^i(t), DV_\nu(t, \tilde{Q}^i(t)))) dt + \sigma B^i(\cdot) \right)(t). \quad (4.6)
\]

Once again using standard martingale arguments it follows that for \(0 \leq s \leq t \leq T\), \(B^i(t) - B^i(s)\) is independent of \(\sigma\{(\tilde{Q}^i(u), \tilde{R}^i(u), \tilde{Y}^i(u), B^i(u)) : u \leq s\}\). From weak uniqueness property noted in Section 3.2 and the fact that \(\tilde{\nu}\) is a fixed point of \(\Phi\) we now have that \(\tilde{Q}^i\) has distribution \(\tilde{\nu}\) for \(i = 2, 3, \cdots\). Using the fact that \(\{B^i\}\) are mutually independent, a simple argument based on Girsanov’s theorem shows that \(\{Q^i\}\) are mutually independent as well. This characterize the distribution of \(\{\tilde{Q}^i\}_{i \geq 2}\) as \(\tilde{\nu}^{\otimes \infty}\). We now have from (4.6), the definition of \(\Xi\), and the law of large numbers that

\[
\lim_{n \to \infty} \tilde{\nu}^n_{-i} = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^n \delta_{\tilde{Q}^n,i} = \lim_{n \to \infty} \frac{1}{n-1} \sum_{i=2}^n \delta_{\tilde{Q}^n,i} = \lim_{m \to \infty} \frac{1}{m-1} \sum_{i=2}^m \delta_{\tilde{Q}^i} = \tilde{\nu}.
\]

The result follows. \(\square\)

**Remark 4.1** The above proof also shows that if \(\hat{\alpha}^{-1}_{-1} = \hat{\alpha}^n\), namely \(\beta^n = \hat{\alpha}^{n,1}\), then (4.6) holds for all \(i \geq 1\) and the law of \(\{\tilde{Q}^i\}_{i \geq 1}\) is \(\tilde{\nu}^{\otimes \infty}\).

### 4.2 Same strategy for all players

In this section we prove the equality in (4.3).

**Proposition 4.2** Suppose Assumptions 3.1 and 4.1 are satisfied. Let \(\hat{\alpha}^n = \{\hat{\alpha}^{n,i}\}_{i=1}^\infty\) be as in (4.1). Then for all \(i \geq 1\)

\[
\lim_{n \to \infty} J^{n,i}(\tilde{Q}^n(0); \hat{\alpha}^n) = V_\nu(0, x). \quad (4.7)
\]

**Proof:** Without loss of generality we assume \(i = 1\). From the proof of Proposition 4.1 (see Remark 4.1)

\[
(\tilde{Q}^{n,1}, \tilde{Y}^{n,1}, \tilde{R}^{n,1}) \Rightarrow (\tilde{Q}^1, \tilde{Y}^1, \tilde{R}^1)
\]

where the processes on the right side are given through (4.6) with \(i = 1\). Let

\[
\hat{\alpha}(t, \tilde{\nu}(t), \tilde{Q}^1(t), DV_\nu(t, \tilde{Q}^1(t))) = \gamma(t, \tilde{Q}^1(t)), \ t \in [0, T].
\]

23
Recall that $f$ and $g$ are bounded continuous functions and from Proposition 4.1 we have that, for every $t \in [0, T]$,
\[(\tilde{v}^n(t), \tilde{Q}^{n,1}(t), \alpha^{n,1}(t)) \Rightarrow (\tilde{v}(t), \tilde{Q}^1(t), \gamma(t, \tilde{Q}^1(t))).\]

This shows that
\[
\mathbb{E} \left[ \int_0^T f(t, \tilde{v}^n(t), \tilde{Q}^{n,1}(t), \alpha^{n,1}(t)) dt + g(\tilde{v}^n(T), \tilde{Q}^{n,1}(T)) \right]
\rightarrow \mathbb{E} \left[ \int_0^T f(t, \tilde{v}(t), \tilde{Q}^1(t), \gamma(t, \tilde{Q}^1(t))) dt + g(\tilde{v}(T), \tilde{Q}^1(T)) \right].
\] (4.8)

Also by continuity of $y$ and $r$
\[
\left(y(\cdot, \tilde{v}^n(\cdot)), r(\cdot, \tilde{v}^n(\cdot)), \tilde{Y}^{n,1}(\cdot), \tilde{R}^{n,1}(\cdot)\right) \Rightarrow \left(y(\cdot, \tilde{v}(\cdot)), r(\cdot, \tilde{v}(\cdot)), \tilde{Y}^1(\cdot), \tilde{R}^1(\cdot)\right)
\]
in $D([0, T] : \mathbb{R}^4)$. It then follows (cf. [24] Lemma 2.4)
\[
\left(\int_0^T y(t, \tilde{v}^n(t)) d\tilde{Y}^{n,1}(t), \int_0^T r(t, \tilde{v}^n(t)) d\tilde{R}^{n,1}(t)\right) \Rightarrow \left(\int_0^T y(t, \tilde{v}(t)) d\tilde{Y}^1(t), \int_0^T r(t, \tilde{v}(t)) d\tilde{R}^1(t)\right)
\] (4.9)
as $n \to \infty$. Also from Lemma 2.1
\[
\mathbb{E} \left[ (\tilde{Y}^{n,1}(T))^2 + (\tilde{R}^{n,1}(T))^2 \right] \leq c_S^2 \mathbb{E} \left[ \sup_{0 \leq s \leq T} |\tilde{\zeta}^{n,1}(s)|^2 \right].
\] (4.10)

Next note that
\[
\mathbb{E} \sup_{0 \leq t \leq T} (\tilde{A}^{n,1}(t))^2 \leq 4 \frac{1}{e_n} \int_0^T \lambda^{n,1}(s) ds \leq 4C_0T,
\]
where $C_0$ is as in (4.4). Similarly, $\mathbb{E} \sup_{0 \leq t \leq T} (\tilde{D}^{n,1}(t))^2 \leq 4C_0T$. Combining these estimates
\[
\sup_n \mathbb{E} \left[ \sup_{0 \leq t \leq T} |\tilde{\zeta}^{n,1}(s)|^2 \right] < \infty
\]
which combined with (4.9) and the boundedness of $y, r$ implies
\[
\sup_n \mathbb{E} \left[ \int_0^T y(t, \tilde{v}^n(t)) d\tilde{Y}^{n,1}(t) + \int_0^T r(t, \tilde{v}^n(t)) d\tilde{R}^{n,1}(t) \right]^2 < \infty.
\]

Combining this with the weak convergence in (4.9) we have
\[
\mathbb{E} \left[ \int_0^T y(t, \tilde{v}^n(t)) d\tilde{Y}^{n,1}(t) + \int_0^T r(t, \tilde{v}^n(t)) d\tilde{R}^{n,1}(t) \right]
\rightarrow \mathbb{E} \left[ \int_0^T y(t, \tilde{v}(t)) d\tilde{Y}^1(t) + \int_0^T r(t, \tilde{v}(t)) d\tilde{R}^1(t) \right].
\] (4.11)
Combining (4.8) and (4.11) and recalling from the optimality of \( \hat{\alpha} \) that

\[
V_\nu(0, x) = \mathbb{E} \left[ \int_0^T f(t, \tilde{\nu}(t), \tilde{Q}^1(t), \hat{\alpha}(t, \tilde{\nu}(t)), \tilde{Q}^1(t), D\nu(t, \tilde{Q}^1(t))) dt + g(\tilde{\nu}(T), \tilde{Q}^1(T)) \right. \\
+ \left. \int_0^T g(t, \tilde{\nu}(t)) d\tilde{Y}^1(t) + \int_0^T r(t, \tilde{\nu}(t)) d\tilde{R}^1(t) \right],
\]

we have the desired convergence \( \lim_{n \to \infty} J_{\nu,1}^n(\tilde{Q}^n(0); \hat{\alpha}^n) = V_\nu(0, x) \). \( \square \)

4.3 Deviation of Player 1

In this section we prove the inequality on the right side of (4.3).

**Proposition 4.3** Suppose Assumptions 3.1 and 4.1 hold. Let \( \tilde{\nu}, \beta^n, \tilde{\alpha}^n_i \) be as introduced at the beginning of Section 4. Then for each \( i \geq 1 \)

\[
\lim inf_{n \to \infty} J_{\nu,1}^n(\tilde{Q}^n(0); \tilde{\alpha}^n_i) \geq V_\nu(0, x).
\]

**Proof:** As before, we assume without loss of generality that \( i = 1 \). We will need to argue the tightness of the control sequence \( \{\beta^n\} \) in an appropriate space. For this it will be convenient to consider a relaxed control formulation. Consider the relaxation of the stochastic control problem in (3.1)-(3.3) where the control space \( U \) is replaced by \( \mathcal{P}(U) \), the drift function \( b \) is replaced by the function \( b_{\mathcal{R}} : [0, T] \times [0, L] \times \mathcal{P}(U) \rightarrow \mathbb{R} \) defined as

\[
b_{\mathcal{R}}(t, x, r) = \int_U b(t, x, u)r(du), \quad (t, x, r) \in [0, T] \times [0, L] \times \mathcal{P}(U),
\]
and the running cost function \( f \) is replaced by \( f_{\mathcal{R}} : [0, T] \times \mathcal{P}([0, L]) \times [0, L] \times \mathcal{P}(U) \rightarrow \mathbb{R} \), defined as

\[
f_{\mathcal{R}}(t, \eta, x, r) = \int_U f(t, \eta, x, u)r(du), \quad (t, \eta, x, r) \in [0, T] \times \mathcal{P}([0, L]) \times [0, L] \times \mathcal{P}(U).
\]

Also, the class of admissible controls \( \mathcal{A}(\Xi, t, x, \nu) \) is replaced by \( \mathcal{A}_{\mathcal{R}}(\Xi, t, x, \tilde{\nu}) \) of pairs \( (\alpha_{\mathcal{R}}, Z) \) that are similar to pairs \( (\alpha, Z) \) introduced above (3.1) except that \( \alpha_{\mathcal{R}} \in \mathcal{P}(U) \) valued rather than \( U \) valued and (3.1) holds with \( \tilde{b}(u) = b(u, X(u), \alpha(u)) \) replaced with \( b_{\mathcal{R}}(u, X(u), \alpha_{\mathcal{R}}(u)) \). The corresponding cost function \( J_{\rho, \mathcal{R}} \) is defined by (3.2) with \( f \) replaced by \( f_{\mathcal{R}} \). The value function in this relaxed formulation, denoted as \( V_{\nu, \mathcal{R}} \), is given by (3.3) with \( \mathcal{A} \) replaced by \( \mathcal{A}_{\mathcal{R}} \). Define the function \( h_{\mathcal{R}} \) by (3.7), replacing \( (f, b) \) with \( (f_{\mathcal{R}}, b_{\mathcal{R}}) \). Then, from Assumption 3.1(b),

\[
H(t, \eta, x, p) = \inf_{u \in U} h(t, \eta, x, u, p) = \inf_{r \in \mathcal{P}(t)} h_{\mathcal{R}}(t, \eta, x, r, p).
\]

This shows that \( V_\nu \) and \( V_{\nu, \mathcal{R}} \) are both solutions of the PDE (3.5) - (3.6). In view of the uniqueness result from Proposition 3.4 \( V_\nu = V_{\nu, \mathcal{R}} \).

Let \( \beta^n_{\mathcal{R}}(t) = \delta_{\beta^n(t)} \), \( t \in [0, T] \) and define \( \tilde{\beta}^n_{\mathcal{R}} \in \mathcal{M}(U \times [0, T]) \) as

\[
\tilde{\beta}^n_{\mathcal{R}}(du, dt) = \beta^n_{\mathcal{R}}(t)(du)dt,
\]
where \( \mathcal{M}(U \times [0,T]) \) is the space of finite measures on \( U \times [0,T] \) equipped with the topology of weak convergence. Then we can rewrite

\[
J^{n,1}(\tilde{Q}^n(0), \alpha^{n,1}) = \mathbb{E} \left[ \int_{U \times [0,T]} f(t, \tilde{v}^n(t), \tilde{Q}^{n,1}(t), u) \tilde{\beta}_R^n(du \, dt) + g(\tilde{v}^n(T), \tilde{Q}^{n,1}(T)) \right]
\]

\[
+ \int_0^T y(t, \tilde{v}^n(t)) \, d\tilde{Y}^{n,1}(t) + \int_0^T r(t, \tilde{v}^n(t)) \, d\tilde{R}^{n,1}(t)
\]

and

\[
(\tilde{Q}^{n,1}, \tilde{Y}^{n,1}, \tilde{R}^{n,1}) = \Gamma \left( \tilde{Q}^{n,1}(0) + \int_{U \times [0,T]} b(s, \tilde{Q}^{n,1}(s), u) \tilde{\beta}^n_R(du \, ds) + \tilde{A}^{n,i} - \tilde{D}^{n,i} + o(1) \right).
\]

From Proposition 4.1, \( \tilde{v}^n \Rightarrow \tilde{v} \). Also, the arguments of the same proposition show that

\[
\left\{ \tilde{Q}^{n,1}(-), \tilde{A}^{n,i} - \tilde{D}^{n,1}(-), \int_{U \times [0,T]} b(s, \tilde{Q}^{n,1}(s), u) \tilde{\beta}^n_R(du \, ds) \right\}_{n \in \mathbb{N}}
\]

are \( C \)-tight in \( \mathcal{D}([0,T] : \mathbb{R}^3) \). Furthermore, since \( U \times [0,T] \) is compact and \( \tilde{\beta}^n_R(U \times [0,T]) = T \), the sequence \( \{\tilde{\beta}^n_R\}_{n \in \mathbb{N}} \) is tight in \( \mathcal{M}(U \times [0,T]) \). Suppose now that along a subsequence (labeled once more as \( \{n\} \)) the sequence in (4.11) along with \( \{\tilde{\beta}^n_R\} \) converges in distribution to \((\tilde{Q}^1, \sigma B^1, \tilde{\nu}, \tilde{\beta}_R)\). Then from the Lipschitz property of \( b \) (Assumption 3.1(a)) we have, for \( t \in [0,T] \),

\[
\tilde{\nu}(t) = \int_{U \times [0,T]} b(s, \tilde{Q}^1(s), u) \tilde{\beta}_R(du \, ds)
\]

\[
= \int_0^t b_R(s, \tilde{Q}^1(s), \tilde{\beta}_R(s)) \, ds,
\]

where \( \beta_R(s) \) is obtained by disintegrating \( \bar{\beta} \), i.e. \( \beta_R(du \, ds) = \beta_R(s)(du) \, ds \). Also as in the proof of Proposition 4.1, it can be argued that \( B^1 \) is a standard Brownian motion and thus we can conclude as in the proof of (4.10)

\[
\tilde{Z}(t) \equiv (\tilde{Q}^1, \tilde{Y}^1, \tilde{R}^1)(t) = \Gamma \left( x + \int_0^T b_R(t, \tilde{Q}^1(t), \tilde{\beta}_R(t)) \, dt + \sigma B^1(\cdot) \right)(t), \quad t \in [0,T].
\]

Once again, by a standard martingale argument, one can argue that for \( 0 \leq s \leq t \leq T \), \( B^1(t) - B^1(s) \) is independent of

\[
\{ \tilde{Q}^1(s'), \tilde{Y}^1(s'), \tilde{R}^1(s'), \tilde{\beta}_R(A \times [0,s']) : 0 \leq s' \leq s \} \quad \text{for } A \in \mathcal{B}(U) \}
\]

Thus denoting by \((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\nu})\) the probability space on which the limit processes are defined, \( \Xi = (\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}, \tilde{\nu}, B^1) \) is a system and \((\tilde{\beta}_R, \tilde{Z}) \in \mathcal{A}_R(\Xi, 0, x, \tilde{\nu}) \). Exactly as in the proof of Proposition 4.2 we see that the convergence in (4.11) holds. Also using the weak convergence

\[
(\tilde{v}^n, \tilde{Q}^{n,1}, \tilde{\beta}^n_R) \Rightarrow (\tilde{\nu}, \tilde{Q}^1, \tilde{\beta}_R)
\]
and the Lipschitz property of $f$ in Assumption 3.1 we have

$$
E \left[ \int_{U \times [0,T]} f(t, \nu^n(t), \hat{Q}^{n,1}(t), u) \beta^n_R(du \, dt) + g(\nu^n(T), \hat{Q}^{n,1}(T)) \right] \\
\rightarrow E \left[ \int_0^T f(t, \nu(t), \hat{Q}^1(t), u) \beta_R(du \, dt) + g(\nu(T), \hat{Q}^1(T)) \right].
$$

(4.18)

Combining the above convergence properties, we have as $n \to \infty$,

$$
J^{n,1}(\hat{Q}^n(0), \tilde{\alpha}^{n,-1}) \rightarrow J(\nu, R(0, x, \hat{Z}) \geq V(\nu, R(0, x)) = V(0, x).
$$

Since the above holds for an arbitrary convergent subsequence of processes in (4.16) and the sequence $\{\beta_R^n\}$, the result follows.

\[\square\]

5 Numerical scheme for the MFG

Construction of an asymptotic Nash equilibrium for the queuing system in Section 2.1 requires the solution of the MFG introduced in Section 3. In general, tractable closed form expressions for the MFG solution are not available and thus one needs suitable numerical approximations. In this section we will use the Markov chain approximation method ([45]) to construct numerical solutions of the MFG. The main result of the section is Theorem 5.1. The numerical scheme is composed of two main steps. First, in Section 5.1 we construct a finite state, discrete time, controlled Markov chain and provide a numerical scheme to construct a measure over $C([0,T] : [0,L])$. Then, in Section 5.2 we show that, under assumptions that include the existence of a unique solution of the MFG, the measure constructed from the chain converges to the solution of the MFG over a small time interval. Throughout the section we assume that Assumption 3.1 is satisfied and that (3.16) holds. We will introduce additional assumptions as needed.

We now introduce the controlled Markov chain that will be used to approximate the solution of the MFG.

### 5.1 Approximating controlled Markov chains

In this section we construct a discrete-time Markov decision problem (MDP) that is used in Section 5.2 to approximate the MFG from Section 3.

Fix a discretization parameter $h > 0$ such that $L$ is an integer multiple of $h$. Denote the $h$-grid $\{-h, 0, h, \ldots, L + h\}$ by $S^h$. This is a discretized version of the state space $[0, L]$. Since 0 and $L$ are reflecting barriers for the state process $X$ in Section 3.1, we will consider two types of transition steps for the approximating chain. The first, which occurs when the chain is away from the boundary, will be referred to as the rate control step and the second occurs at the end points $L + h$ and $-h$ and is referred to as the reflection step.

**Rate control step.** For every $t \in [0, T]$, $\eta \in \mathcal{P}([0,L])$, $u \in U$, and $x \in S^h_0 \doteq S^h \setminus \{-h, L + h\}$
let
\[ q^h(t, \eta, u; x, x \pm h) = \pm hb(t, \eta, x, u) + \frac{\sigma^2}{2\sigma^2}, \]
(5.1)

Note that
\[ \sum_{y \in \{x \pm h\}} q^h(t, \eta, u; x, y) = 1 \]
and that for \(0 < h < \sigma^2/cB\), the transition probabilities are positive. Hereafter, these inequalities on \(h\) are in force. Also, define
\[ \Delta^h \doteq \frac{h^2}{\sigma^2}. \]

This will be used to define the continuous time interpolation of the controlled Markov chain. Denote the \(\Delta^h\)-grid \(\{0, \Delta^h, 2\Delta^h, \ldots, T - \Delta^h\}\) by \(T^h\). Note that \(\Delta^h \to 0\) as \(h \to 0\). One can verify that the following local consistency conditions (cf. [45]) hold for every \(x \in S^h\).
\[ m^h_0(t, \eta, u, x) = \int_{S^h} (\tilde{x} - x)q^h(t, \eta, u; x, d\tilde{x}) = b(t, \eta, x, u)\Delta^h, \]
(5.2)
\[ (\sigma^h_0)^2(t, \eta, u, x) = \int_{S^h} (\tilde{x} - x - m^h_0)^2q^h(t, \eta, u; x, d\tilde{x}) = \sigma^2\Delta^h - (b(t, \eta, u, x)\Delta^h)^2. \]
(5.3)

**Reflection step.** Such a step occurs only when \(x \in \{-h, L+h\}\). For every \(t \in [0, T]\), \(\eta \in \mathcal{P}([0, L])\), and \(u \in U\) let
\[ q^h(t, \eta, u; L + h, L) = q^h(t, \eta, u; -h, 0) = 1. \]

We will now define a controlled Markov chain \(\{X^{h,\nu}_n\}_{n \in \mathbb{N}_0}\) associated with the parameter \(h\) and a measure \(\nu \in \mathcal{P}_{T,L}\). We will assume that \(T/\Delta^h = I(h)\) and \(L/h\) are integers.

**Construction 5.1**

- Define \(X^{h,\nu}_0 = x^h_0\), \(t^{h,\nu}_0 = 0\) and let \(\alpha^{h,\nu}_{-1}\) be a fixed element of \(U\).

- Having defined for \(i = 0, 1, \ldots, n\) time instants \(t^{h,\nu}_i < T\) and random variables \(X^{h,\nu}_i, \alpha^{h,\nu}_{-1}\) with values in \(S^h\) and \(U\) respectively, let \(\mathcal{F}^{h,\nu}_i = \sigma\{X^{h,\nu}_j, \alpha^{h,\nu}_{j-1} : j = 0, 1, \ldots, i\}\).

- Choose the control \(\alpha^{h,\nu}_n\) for the \(n\)-th step that is a \(U\) valued \(\mathcal{F}^{h,\nu}_n\) measurable random variable and let \(X^{h,\nu}_{n+1}\) be such that the conditional distribution of \(X^{h,\nu}_{n+1}\) given \(\mathcal{F}^{h,\nu}_n\) is \(q^h(t^{h,\nu}_n, \nu(t^{h,\nu}_n), \alpha^{h,\nu}_n, X^{h,\nu}_n, \cdot)\) where \(\nu(t)\) denotes the marginal distribution of \(\nu\) at time instant \(t\). Also define
\[ t^{h,\nu}_{n+1} = t^{h,\nu}_n + \Delta^h 1\{X^{h,\nu}_n \notin \{-h, L+h\}\} \]
(5.4)
where the indicator in the above definition will ensure that when we do a continuous time interpolation of the chain, reflection steps ‘occur instantaneously’.
If $\alpha_{n}^{h,\nu} = \vartheta(t_{n}^{h,\nu}, X_{n}^{h,\nu})$ for some $\vartheta : \mathbb{T}^{h} \times S_{0}^{h} \to U$ then the function $\vartheta$ is referred to as a feedback control.

**Some auxiliary processes.** We will now introduce some processes that will be useful in the analysis of the $h$-th Markov chain. Consider the piecewise constant processes

$$
(X_{h,\nu}^{h,\nu}(t), \alpha_{n}^{h,\nu}(t)) = (X_{n}^{h,\nu}(t), \alpha_{n}^{h,\nu}(t)), t \in [0, T],
$$

(5.5)

where

$$
n_{h,\nu}(t) = \max\{n : t_{n}^{h,\nu} = jh\}, t \in [j\Delta^{h}, (j+1)\Delta^{h}), j = 0, \ldots, I(h) - 1.
$$

(5.6)

Let $F_{h,\nu}^{h,\nu}(0) = B_{h,\nu}^{h,\nu}(0) = Y_{h,\nu}^{h,\nu}(0) = R_{h,\nu}^{h,\nu}(0) = 0$ and for every $t \in [0, T]$ let

$$
F_{h,\nu}^{h,\nu}(t) = \sum_{j=\max\{0\}}^{n_{h,\nu}(t)-1} \mathbb{E}\left[ X_{j+1}^{h,\nu} - X_{j}^{h,\nu} \mid F_{j}^{h,\nu} \right] 1\{X_{j}^{h,\nu} \notin \{-h, L+h\}\},
$$

(5.7)

$$
B_{h,\nu}^{h,\nu}(t) = \frac{1}{\sigma} \sum_{j=\max\{0\}}^{n_{h,\nu}(t)-1} \left( X_{j+1}^{h,\nu} - X_{j}^{h,\nu} - \mathbb{E}\left[ X_{j+1}^{h,\nu} - X_{j}^{h,\nu} \mid F_{j}^{h,\nu} \right] \right) 1\{X_{j}^{h,\nu} \notin \{-h, L+h\}\},
$$

$$
Y_{h,\nu}^{h,\nu}(t) = \sum_{j=\max\{0\}}^{n_{h,\nu}(t)-1} (X_{j+1}^{h,\nu} - X_{j}^{h,\nu}) 1\{X_{j}^{h,\nu} = -h\} = h \sum_{j=\max\{0\}}^{n_{h,\nu}(t)-1} 1\{X_{j}^{h,\nu} = -h\},
$$

$$
R_{h,\nu}^{h,\nu}(t) = \sum_{j=\max\{0\}}^{n_{h,\nu}(t)-1} (X_{j}^{h,\nu} - X_{j+1}^{h,\nu}) 1\{X_{j}^{h,\nu} = L+h\} = h \sum_{j=\max\{0\}}^{n_{h,\nu}(t)-1} 1\{X_{j}^{h,\nu} = L+h\}.
$$

One can verify that the following representation holds

$$
(X_{h,\nu}^{h,\nu}, Y_{h,\nu}^{h,\nu}, R_{h,\nu}^{h,\nu})(t) = \Gamma \left( \sigma_{0}^{h} + F_{h,\nu}^{h,\nu}(\cdot) + \sigma B_{h,\nu}^{h,\nu}(\cdot) \right)(t), \ t \in [0, T].
$$

(5.8)

Also, from (5.2) and (5.3) it follows that, on the set \{X_{n}^{h,\nu} \notin \{-h, L+h\}\}

$$
\mathbb{E}\left[ X_{n+1}^{h,\nu} - X_{n}^{h,\nu} \mid F_{n}^{h,\nu} \right] = b(t_{n}^{h,\nu}, \nu(t_{n}^{h,\nu}), X_{n}^{h,\nu}(t_{n}^{h,\nu}), \alpha_{n}^{h,\nu}(t_{n}^{h,\nu}))\Delta^{h,\nu},
$$

$$
\mathbb{E}\left[ \left( X_{n+1}^{h,\nu} - X_{n}^{h,\nu} - \mathbb{E}\left[ X_{n+1}^{h,\nu} - X_{n}^{h,\nu} \mid F_{n}^{h,\nu} \right] \right)^{2} \mid F_{n}^{h,\nu} \right] = \sigma^{2}\Delta^{h,\nu} + o(\Delta^{h,\nu}).
$$

Define $\mathcal{G}_{t}^{h,\nu} = \mathcal{F}_{n}^{h,\nu}(t)$. Then since $n_{h,\nu}(t)$ for each fixed $t$ is a $\{\mathcal{F}_{j}^{h,\nu}\}$ stopping time, we have by optional sampling theorem that $B_{h,\nu}^{h,\nu}(\cdot)$ is a $\{\mathcal{G}_{t}^{h,\nu}\}$ martingale. Also, from the above,

$$
F_{h,\nu}^{h,\nu}(t) = \int_{0}^{t} b(t_{s}^{h,\nu}, \nu(t_{s}^{h,\nu}), X_{s}^{h,\nu}(s), \alpha_{s}^{h,\nu}(s)) ds
$$

(5.9)

where

$$
l_{s}^{h} = [s/\Delta^{h}]\Delta^{h}, s \in [0, T].
$$
Cost function for the MDP. For every \((t,x)\in \mathbb{T}^h \times \mathbb{S}^h\) and any admissible control \(\alpha^{h,\nu}\) used to construct the \(h\)-th controlled Markov chain, define the associated cost

\[
J^{h,\nu}(t,x,\alpha^{h,\nu}) \doteq \mathbb{E} \left[ \int_t^T f(l^h(s),\nu(l^h(s)),X^{h,\nu}(s),\alpha^{h,\nu}(s))ds + g(\nu(T),X^{h,\nu}(T)) \right. \\
+ \left. \int_t^T y(s)dY^{h,\nu}(s) + \int_t^T r(s)dR^{h,\nu}(s) \mid X^{h,\nu}(t) = x \right].
\]

The value function associated with the above cost is given by,

\[
V^h_\nu(t,x) \doteq \inf_{\alpha} J^{h,\nu}(t,x,\alpha),
\]
where the infimum is taken over all admissible controls.

We now provide properties of the value function \(V^h_\nu\) and the optimal strategy in the \(h\)-th MDP. For every \((t,x,\nu)\in \mathbb{T}^h \times \mathbb{S}^h \times \mathcal{P}_{T,L}\), define the \(h\)-th finite difference of the value function w.r.t. \(x\), as follows

\[
D^h_x V^h_\nu(t,x) \doteq \frac{1}{2h}(V^h_\nu(t + \Delta^h,x + h) - V^h_\nu(t + \Delta^h,x - h)),
\]
where we use the convention that

\[
V^h_\nu(t + \Delta^h, L + h) \doteq r(t + \Delta^h)h + V^h_\nu(t + \Delta^h, L),
\]
\[
V^h_\nu(t + \Delta^h, -h) \doteq y(t + \Delta^h)h + V^h_\nu(t + \Delta^h, 0).
\]

**Lemma 5.1** The optimal control in the \(h\)-th MDP is given in state feedback form as

\[
\dot{\varrho}^{h,\nu}(t,x) = \hat{\alpha}(t,\nu(t),x,D^h_x V^h_\nu(t,x)), \quad (t,x) \in \mathbb{T}^h \times \mathbb{S}^h.
\]

Letting

\[
\alpha^{h,\nu}(t,x) \doteq \varrho^{h,\nu}(l^h(t),x), \quad (t,x) \in [0,T] \times \mathbb{S}^h,
\]
there exists a constant \(c_d(T) \in (0,\infty)\), such that for every \((t,\nu)\in \mathbb{T}^h \times \mathcal{P}_{T,L}\) and for every \(h\), one has,

\[
V^h_\nu(t,X^{h,\nu}(t)) + \sigma \sum_{s \in \mathbb{T}^h, s \geq t} D^h_x V^h_\nu(s,X^{h,\nu}(s))(B^{h,\nu}(s + \Delta^h) - B^{h,\nu}(s))
\]
\[
= g(\nu(T),X^{h,\nu}(T)) + \int_t^T f(l^h(s),\nu(l^h(s)),X^{h,\nu}(s),\alpha^{h,\nu}(s,X^{h,\nu}(s)))ds
\]
\[
+ \int_t^T y(s)dY^{h,\nu}(s) + \int_t^T r(s)dR^{h,\nu}(s),
\]
where \((X^{h,\nu},B^{h,\nu},Y^{h,\nu},R^{h,\nu})\) are as in \((5.5),(5.7)\) with \(\alpha^{h,\nu}_n\) replaced with the optimal feedback control \(\varrho^{h,\nu}\) and for all \((t,x)\in \mathbb{T}^h \times \mathbb{S}^h\)

\[
|D^h_x V^h_\nu(t,X^{h,\nu}(t))| \leq c_d(T).
\]

Finally, \(c_d(T)\) can be taken to be a monotonically increasing function of \(T\).
The proof of the lemma is given in Section 6.

The induced measure \( \Phi^h(\nu) \). Recall from (5.6) that for \( j = 0,1,\ldots I(h) \), \( h_{ij} = \max\{i : t_{ij} = j\Delta^h\} \). Let \( \{X^h(\nu)(t)\}_{t \in [0,T]} \) be the continuous stochastic process which is linear on \([j\Delta^h,(j+1)\Delta^h]\) and equals \( X_{ij}^h \) at \( t = j\Delta^h \), for \( j = 0,\ldots I(h) - 1 \), where \( \{X_{ij}^h\} \) is the controlled Markov chain constructed using the optimal feedback control \( \hat{\alpha}_{ij} \). Let

\[
\Phi^h(\nu) = \mathbb{P} \circ (\hat{X}^h)^{-1}.
\]  
(5.17)

Next, we show that, under suitable conditions, \( \Phi^h \) is a contraction up to an \( O(h^2) \) term, over a small time interval. This ‘almost-contraction’ property lies at the heart of our main result, Theorem 5.1.

An almost-contraction property. Recall that we assume that Assumption 3.1 is satisfied. In addition, we will make the following assumption on a Lipschitz property of the function \( \hat{\alpha} \) from (3.9).

Assumption 5.1. There exists \( c_\alpha \in (0,\infty) \) such that for every \( t \in [0,T] \), \( \eta,\eta' \in \mathcal{P}([0,L]) \), \( x,x' \in [0,L] \), and \( p,p' \in \mathbb{R} \),

\[
|\hat{\alpha}(t,\eta,p) - \hat{\alpha}(t,\eta',p')| \leq c_\alpha(W(\eta,\eta') + |x - x'| + |p - p'|).
\]  
(5.18)

The following lemma gives a sufficient condition for Assumption 5.1 to hold.

Lemma 5.2. Suppose that the drift and the cost functions satisfy the following properties.

(a) For every \( (t,\eta,x,\alpha) \in [0,T] \times \mathcal{P}([0,L]) \times [0,L] \times U \),

\[
b(t,\eta,x,\alpha) = b_1(t,\eta,x) + b_2(t)\alpha.
\]  
(5.19)

(b) There exists \( c_m \in (0,\infty) \) such that for every \( t \in [0,T] \), \( \eta,\eta' \in \mathcal{P}([0,L]) \), \( x,x' \in [0,L] \), and \( \alpha,\alpha' \in U \),

\[
f(t,\eta,x,\alpha') - f(t,\eta,x,\alpha) - (\alpha' - \alpha)f_\alpha(t,\eta,x,\alpha) \geq c_m(\alpha' - \alpha)^2.
\]  
(5.20)

(c) The map \( \alpha \mapsto f(t,\eta,x,\alpha,p) \) is continuously differentiable for every \( (t,\eta,x,p) \in [0,T] \times \mathcal{P}([0,L]) \times [0,L] \times \mathbb{R} \) and there exists \( c_\ell \in (0,\infty) \) such that for every \( t \in [0,T] \), \( \eta,\eta' \in \mathcal{P}([0,L]) \), \( x,x' \in [0,L] \), \( \alpha \in U \), and \( p,p' \in \mathbb{R} \),

\[
|f_\alpha(t,\eta,x,\alpha,p) - f_\alpha(t,\eta',x',\alpha,p')| \leq c_\ell(W(\eta,\eta') + |x - x'| + |p - p'|).
\]  
(5.21)

Then Assumption 5.1 is satisfied.

The proof of the lemma is deferred to Section 6.

Remark 5.1. The conditions in the above lemma are not new to the literature of MFG. For example, Assumptions (A.1), (A.2) and (A.3) in [21] are stronger. Also, parts (b) and (c) above that concern the running cost are imposed by [27], which studies a rate control problem.
Proposition 5.1 Suppose that Assumption 5.1 is satisfied. Then there exist \( \hat{T} > 0 \), \( \hat{h} > 0 \) and \( q \in (0,1) \), such that for every \( T \leq \hat{T} \), \( \nu, \nu' \in \mathcal{P}_{T,L} \) and \( h \in (0, \hat{h} \wedge \hat{T}) \),

\[
W_1^2(\Phi^h(\nu), \Phi^h(\nu')) \leq q\left(h^2 + W_1^2(\nu, \nu')\right).
\]

Proof of the proposition is given in Section 6.

5.2 Approximating the solution of the MFG.

We now provide the numerical scheme that approximates the solution of the MFG from Section 3.

Construction 5.2 Let \( \hat{T}, \hat{h} \) be as in Proposition 5.1. Fix \( T < \hat{T} \) and \( (x_0^h, \nu^h, h) \in \mathcal{S}^h \times \mathcal{P}_{T,L} \times (0, \hat{h} \wedge \hat{T}) \). Let \( \{X_n^{h,\nu^h}\} \) be the \( h \)-th Markov chain from Construction 5.1 associated with the optimal control \( \hat{\alpha}^{h,\nu^h} \). Having defined for \( m \in \mathbb{N} \) the process \( \{X_n^{h,\nu^m}\} \), set \( \nu^{m+1} = \Phi^h(\nu^m) \) and let \( \{X_n^{h,\nu^{m+1}}\} \) be the \( h \)-th Markov chain from Construction 5.1 associated with the optimal control \( \hat{\alpha}^{h,\nu^{m+1}} \).

With \( q \in (0,1) \) as in Proposition 5.1 we get that for every \( h \) as in Construction 5.2 and every \( m \in \mathbb{N} \),

\[
W_1^2(\Phi^h(\nu^k), \nu^k) = W_1^2(\Phi^h(\nu^k), \Phi^h(\nu^{k-1})) \leq q(h^2 + W_1^2(\nu^k, \nu^{k-1})).
\]
By iterating this bound we obtain
\[ W^2_1(\Phi^k(\nu^k), \nu^k) \leq \frac{q}{1-q} h^2 + q^{k-1} W^2_1(\nu^2, \nu^1). \] (5.22)

Set
\[ k_h = \min \left\{ k \in \mathbb{N} : W^2_1(\Phi^h(\nu_k), \nu_k) \leq \frac{2q}{1-q} h^2 \right\} \]
and
\[ \nu_h \doteq \nu^{k_h}. \] (5.23)

We note that \( k_h \) depends also on \( \nu^1 \), however, it plays no role in the sequel and is therefore omitted from the notation.

As an immediate consequence of the definition of \( \nu_h \), we get the following proposition, which is key to the proof of the approximation result in Theorem 5.1 below.

**Proposition 5.2** Suppose that Assumption 5.1 is satisfied. Then with \( \hat{T} \) as in Proposition 5.1, for every \( T \leq \hat{T} \)
\[ \lim_{h \to 0} W^2_1(\Phi_h(\nu_h), \nu_h) = 0. \] (5.24)

For the main result of this section (Theorem 5.1), in addition to Assumptions 3.1, 5.1 and the property in (3.16), we also need the following assumption.

**Assumption 5.2** There exists \( x \in [0, L] \) such that \( \lim_{h \to 0} x_0^h = x \). Also, there is a unique \( \bar{\nu} \in \mathcal{P}_{T,L} \) that solves the MFG with initial condition \( x \).

**Remark 5.2** Recall from Section 3.3 that the second part of Assumption 5.2 is satisfied if in addition to Assumption 3.1 Assumption 3.2 holds. Also, from Assumptions 3.1 and 5.2 it follows from arguments of Section 3.2 (see proof of Proposition 3.1) that there is a continuous map \( \gamma : [0, T] \times [0, L] \to U \) such that if there exist a system \( \Xi \) and an \((\alpha_R, Z) \in \mathcal{A}_R(\Xi, 0, x, \nu)\) such that
\[ Z(t) \equiv (X, Y, R)(t) = \Gamma \left( x + \int_0^t b_R(t, \bar{\nu}(t), X(t), \alpha_R(t)) dt + \sigma B(\cdot) \right)(t), \quad t \in [0, T], \] (5.25)
\[ \mathbb{P} \circ X^{-1} = \bar{\nu}, \text{ and} \]
\[ V_\nu(0, x) = J_\nu(0, x, \alpha_R, Z). \] (5.26)

then \( \alpha_R(t, \omega) = \delta_{\gamma(t, X(t, \omega))}, \lambda^0_T \otimes \mathbb{P}-a.s. \)

We can now present the main result of this section. The statement of the theorem and its proof use relaxed controls. Recall the definitions of \( \mathcal{M}(U \times [0, T]) \), \( \mathcal{A}_R \), \( b_R \), and \( J_{\nu, R} \), given in the proof of Proposition 4.3. Define \( \mathcal{M}(U \times [0, T]) \) valued random variable \( \hat{m}_{h,\nu_h} \) as
\[ \hat{m}_{h,\nu_h}(du) ds = \delta_{\alpha_h(\nu_h(s), X^h_{\nu_h(s)})}(du) ds. \]

Also, recall Assumption 3.1 and (3.16) are in force throughout this section.
Theorem 5.1 Suppose that $T \leq \hat{T}$ where $\hat{T}$ is as in Proposition [5.1]. Also suppose that Assumptions 5.1 and 5.2 are satisfied. Consider a sequence $\{h\} \to 0$. Then the sequence

$$(X^{h,\nu_h}, Y^{h,\nu_h}, R^{h,\nu_h}, B^{h,\nu_h}, \hat{m}^{h,\nu_h})$$

introduced in the above paragraph converges in distribution to $(X,Y,R,B,m)$ in $\mathcal{D}([0,T] : \mathbb{R}^4) \times \mathcal{M}(U \times [0,T])$ and

$$\lim_{h \to 0} \nu_h = \nu \quad (5.27)$$

in $\mathcal{P}_{T,L}$ where the limit processes defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ satisfy the following.

(a) $B$ is a $\mathcal{G}_t \doteq \sigma\{B(s), X(s), Y(s), R(s), m(A \times [0,s]) : s \leq t, A \in \mathcal{B}(U)\}$ Brownian motion and so $\Xi \doteq \{\Omega, \mathcal{F}, \{\mathcal{G}_t\}, \mathbb{P}, B\}$ is a system.

(b) Disintegrating $m(du\,ds) = m_s(du)\,ds$, the following relationship holds a.s.

$$Z(t) \doteq (X(t), Y(t), R(t)) = \Gamma \left( x + \int_0^t b(s, \nu(s), X(s), m_s)\,ds + \sigma B(\cdot) \right)(t), \ t \in [0,T].$$

(c) $\mathbb{P} \circ (X)^{-1} = \nu$.

(d) The pair $(m, Z) \in \mathcal{A}_R(\Xi,0,x,\nu)$ and $V_0(0,x) = J_{\nu,R}(0,x,m,Z)$. In particular, with $\gamma$ as in Remark 5.2, $m(du\,ds) = \delta_{\gamma(s,X(s))}(du)\,ds$, and $\nu = \tilde{\nu}$, the unique solution of the MFG.

Proof: Since many steps in the proof are quite standard we will only provide details where appropriate. Using properties [5.2] and [5.3] of the controlled transition probability kernel it can be argued (cf. Proof of [45, Theorem 9.4.1]) that $\{(F^{h,\nu_h}, B^{h,\nu_h})\}_{h>0}$ are tight in $\mathcal{D}([0,T] : \mathbb{R}^2)$. Using this tightness property along with the continuity of the Skorohod map (Lemma 2.1) it now follows that $\{(X^{h,\nu_h}, Y^{h,\nu_h}, R^{h,\nu_h}, B^{h,\nu_h})\}_{h>0}$ is tight in $\mathcal{D}([0,T] : \mathbb{R}^4)$. In fact, this sequence is $C$-tight. Next note that, since $\hat{m}^{h,\nu_h}(U \times [0,T]) = T$ and $U$ is compact, the sequence $\{\hat{m}^{h,\nu_h}\}_{h>0}$ is tight in $\mathcal{M}(U \times [0,T])$. Also, recalling the definition of the interpolated processes (right before (5.17)) it can be checked that

$$\sup_{0 \leq t \leq T} |X^{h,\nu_h}(t) - \hat{X}^{h,\nu_h}(t)| \to 0, \text{ in probability as } h \to 0. \quad (5.28)$$

Combining this with the tightness of $\{X^{h,\nu_h}\}$ and the fact that $\Phi^h(\nu_h) = \mathbb{P} \circ (X^{h,\nu_h})^{-1}$ gives the relative compactness of $\{\Phi^h(\nu_h)\}$ in $\mathcal{P}_{T,L}$. Suppose now that along a subsequence (relabeled again as $\{h\}$)

$$ (X^{h,\nu_h}, Y^{h,\nu_h}, R^{h,\nu_h}, B^{h,\nu_h}, \hat{m}^{h,\nu_h}) \Rightarrow (X,Y,R,B,m) \quad (5.29)$$

and

$$\Phi^h(\nu_h) \to \nu. \quad (5.30)$$

Then by (5.24) we also have that

$$\nu_h \to \nu. \quad (5.31)$$
By standard martingale methods it follows that $B$ is a $\{\mathcal{G}_t\}$ Brownian motion (see e.g. Proof of [45, Theorem 9.4.1]) proving part (a) of the theorem.

Using (5.29), $P^{h,\nu_h}$ converges, along with the above processes, in distribution to

$$
\int_{U \times [0,]} b(s, \nu(s), X(s), u)m(du \, ds) = \int_0 b_\mathcal{R}(s, \nu(s), X(s), m_s) \, ds,
$$

where $m(du \, ds) = m_s(du) \, ds$. Using the continuity property of the Skorohod map we now get (b).

Also, part (c) is immediate on using (5.28), recalling that $\Phi^h(\nu)$ is the probability law of $X_t^{h,\nu_h}$, and by (5.31).

Clearly $(m, Z) \in \mathcal{A}_\mathcal{R}(\Xi, 0, x, \nu)$.

We will now argue that

$$
V_\nu(0, x) = J_{\nu, \mathcal{R}}(0, x, m, Z).
$$

This will prove that $\nu$ solves the MFG and thus from Assumption 5.2 and Remark 5.2 we will get part (d). Since the convergent subsequence was arbitrary, we will get the convergence asserted in the statement of the theorem and complete the proof.

**Proof of (5.32).** For the rest of the proof we consider the subsequence along which the convergence in (5.29) and (5.30) holds. Using arguments similar to those in Proposition 4.2 it can be checked that

$$
\lim_{h \to 0} J^{h,\nu_h}(\hat{\alpha}^{h,\nu_h}) = J_{\nu, \mathcal{R}}(0, x, m, Z),
$$

where we dropped the arguments $t = 0$ and $x$. Next, let $(\beta, Z) \in \mathcal{A}(\bar{\Xi}, 0, x, \nu)$ for some system $\bar{\Xi} = (\bar{\Omega}, \bar{\mathcal{F}}, \{\bar{\mathcal{F}}_t\}, \bar{\mathbb{P}}, \bar{B})$. We will now show that for every $\varepsilon_0 > 0$ there is a sequence of controls $\{\beta^{h,\nu_h}\}$ for the controlled Markov chains defined as in Construction 5.1 with the sequence $\{\nu_h\}$, such that

$$
\limsup_{h \to 0} J^{h,\nu_h}(\beta^{h,\nu_h}) \leq J_{\nu}(0, x, \beta, Z) + \varepsilon_0.
$$

Note that from the optimality property of $\hat{\alpha}^{h,\nu_h}$

$$
\limsup_{h \to 0} J^{h,\nu_h}(\hat{\alpha}^{h,\nu_h}) \leq \limsup_{h \to 0} J^{h,\nu_h}(\beta^{h,\nu_h}).
$$

Combining the above inequality with (5.33) and (5.34) we now get that

$$
V_\nu(0, x) \leq J_{\nu, \mathcal{R}}(0, x, m, Z) = \limsup_{h \to 0} J^{h,\nu_h}(\hat{\alpha}^{h,\nu_h}) \leq \limsup_{h \to 0} J^{h,\nu_h}(\beta^{h,\nu_h}) \leq J_{\nu}(0, x, \beta, Z) + \varepsilon_0.
$$

Since $\varepsilon_0 > 0$, the system $\bar{\Xi}$, and $(\beta, Z) \in \mathcal{A}(\bar{\Xi}, 0, x, \nu)$ are arbitrary we have (5.32), completing the proof of the theorem.

We now prove (5.34). Using arguments as in the proof of [45, Theorem 10.3.1] it can be shown that there is a $\theta_1 : [0, \infty) \to [0, \infty)$ such that $\theta_1(\kappa) \to 0$ as $\kappa \to 0$ and for every $\varepsilon > 0$ there is a system $\bar{\Xi}^\varepsilon = (\bar{\Omega}^\varepsilon, \bar{\mathcal{F}}^\varepsilon, \{\bar{\mathcal{F}}_t^\varepsilon\}, \bar{\mathbb{P}}^\varepsilon, \bar{B}^\varepsilon)$ and $(\beta^\varepsilon, Z^\varepsilon) \in \mathcal{A}(\Xi^\varepsilon, 0, x, \nu)$ with the following properties

- $Z^\varepsilon$ satisfies the following equation for $t \in [0, T]$.

$$
Z^\varepsilon(t) \equiv (X^\varepsilon, Y^\varepsilon, R^\varepsilon)(t) = \Gamma \left( x + \int_0^t b(t, \nu(s), X^\varepsilon(s), \beta^\varepsilon(s)) \, ds + \sigma B^\varepsilon(\cdot) \right)(t). \tag{5.35}
$$
Note that since \( \theta \) theorem. The construction is carried out as follows.

- For some \( \delta > 0 \), \( \beta^\varepsilon \) is piecewise constant on intervals of the form \([l\delta, (l+1)\delta)\), \( l = 0, 1, \ldots, T/\delta \). For some finite set \( U^\varepsilon \subset U \), \( \beta^\varepsilon (s) \) takes values in \( U^\varepsilon \) for every \( s \in [0, T] \).

- For some \( \theta > 0 \), for each \( u \in U^\varepsilon \)

\[
\mathbb{P}^\varepsilon(\beta^\varepsilon(l\delta) = u | B^\varepsilon(s), s \leq l\delta, \beta^\varepsilon(j\delta), j < l) = \mathbb{P}^\varepsilon(\beta^\varepsilon(l\delta) = u | B^\varepsilon(p\theta), p\theta \leq l\delta, \beta^\varepsilon(j\delta), j < l) = F_u^\varepsilon(B^\varepsilon(p\theta), p\theta \leq l\delta, \beta^\varepsilon(j\delta), j < l),
\]

(5.36)

where for suitable \( \hat{I}, \hat{L} \in \mathbb{N} \), \( F_u^\varepsilon : \mathbb{R}^\hat{I} \times (U^\varepsilon)^\hat{L} \to [0, 1] \) is a measurable function such that \( F_u(\cdot, u) \) is continuous on \( \mathbb{R}^\hat{I} \) for every \( u \in (U^\varepsilon)^\hat{L} \).

- Letting \( m^\varepsilon(du) = \delta_{\beta^\varepsilon(t)}(du) dt, m(du) = \delta_{\beta(t)}(du) dt \), as \( \varepsilon \to 0 \),

\[
(X^\varepsilon, Y^\varepsilon, R^\varepsilon, B^\varepsilon, m^\varepsilon) \Rightarrow (X, Y, R, B, m)
\]

in \( \mathcal{D}([0, T] : \mathbb{R}^4) \times \mathcal{M}(U \times [0, T]) \).

- \( J_{\nu}(0, x, \beta^\varepsilon, Z^\varepsilon) \leq J_{\nu}(0, x, \beta, Z) + \theta_1(\varepsilon) \).

We will now use the piecewise constant control \( \beta^\varepsilon \) to construct a collection of control sequences \( \{\beta_{n,\nu}^h\} \) as stated above (5.34) and for which

\[
\lim_{h \to 0} J^h_{\nu}(\beta_{n,\nu}^h, \nu_h) = J_{\nu}(0, x, \beta^\varepsilon, Z^\varepsilon).
\]

(5.37)

Note that since \( \theta_1(\varepsilon) \to 0 \) as \( \varepsilon \to 0 \), this will prove (5.34) and complete the proof of the theorem. The construction is carried out as follows.

- Define \( X_0^{h,\nu_h} = x^{h,\nu_h} \), \( t_0^1 = 0 \) and let \( \beta_{-1,\nu}^h \) be a fixed element of \( U^\varepsilon \).

- Having defined for \( i = 0, 1, \ldots, n \) time instants \( t_i^{h,\nu_h} < T \) and random variables \( X_i^{h,\nu_h}, \beta_{i-1}^{h,\nu_h} \) with values in \( \mathbb{S}^h \) and \( U^\varepsilon \) respectively, let \( F_i^{h,\nu_h} = \sigma\{X_j^{h,\nu_h}, \beta_{j-1}^{h,\nu_h} : j = 0, 1, \ldots i\} \).

- Choose the control \( \beta_{n,\nu}^h \) for the \( n \)-th step that is a \( U^\varepsilon \) valued \( F_n^{h,\nu_h} \) measurable random variable as follows:

  - If \( t_{i-1}^{h,\nu_h} \) and \( t_{i}^{h,\nu_h} \) both lie in \([j\delta, (j+1)\delta)\) for some \( j \), set \( \beta_{n}^{h,\nu_h} = \beta_{i-1}^{h,\nu_h} \).

  - If \( t_{i}^{h,\nu_h} < l\delta \leq t_{i+1}^{h,\nu_h} < (l+1)\delta \) for some \( l \), choose \( \beta_{n}^{h,\nu_h} \) according the the conditional distribution

    \[
    P\left( \beta_{n}^{h,\nu_h} = u \mid X_i^{h,\nu_h}, \beta_{i-1}^{h,\nu_h}, 0 \leq i \leq n \right) = F_u^\varepsilon(B^\varepsilon(p\theta), p\theta \leq l\delta, \beta^\varepsilon(j\delta), j < l),
    \]

    where \( B^\varepsilon(p\theta) \) is defined as in (5.7), \( \beta^\varepsilon(s) = \beta_{n,\nu_h}^{h,\nu_h}(s) \), and \( n^{h,\nu_h}(\cdot) \) is as in (5.6).

- Let \( X_{n+1}^{h,\nu_h} \) be such that its conditional distribution given \( F_n^{h,\nu_h} \) is \( q^h(t_n^{h,\nu_h}, \nu_h(t_n^{h,\nu_h}), \beta_{n}^{h,\nu_h}, X_n^{h,\nu_h}, \cdot) \), where \( \nu_h \) is as introduced above Theorem 5.1. Also define

\[
t_{n+1}^{h,\nu_h} = t_n^{h,\nu_h} + \Delta^h \{X_n^{h,\nu_h} \notin \{L-h, L+h\}\},
\]

36
Now define processes $X^{h,v_h}, F^{h,v_h}, Y^{h,v_h}, R^{h,v_h}$ as in (5.5)–(5.7). Exactly as in the first part of the proof we now have that $\{(X^{h,v_h}, Y^{h,v_h}, R^{h,v_h})\}_h$ is $\mathcal{C}$-tight in $\mathcal{D}([0, T] : \mathbb{R}^4)$ and the sequence $\{m^{h,v_h}\}_{k \geq 1}$, where $m^{h,v_h}(duds) = \delta_{\phi^{h,v_h}(s)}(du)ds$ is tight in $\mathcal{M}(U \times [0, T])$. Arguing as before, if along a further subsequence the convergence (5.29) holds (with $\tilde{m}^{h,v_h}$ replaced with $m^{h,v_h}$) then parts (a) and (b) as in the statement of Theorem 5.1 are satisfied with $\nu$ as in (5.30). Using the continuity property of $F_\nu$ and the fact that the control is piecewise constant with values in a finite set it follows that (cf. Proof of [45, Theorem 10.5.2]) ($B,m$) has the same distribution as $(\nu,F_\nu)$ for every limit point of the chosen further subsequence. Since the chosen further subsequence was arbitrary, this proves the weak convergence of $(Z^{h,v_h}, m^{h,v_h})$ to $(\nu,F_\nu)$ along the subsequence fixed above (5.33) and arguing once again as in the proof of Proposition 4.2 we have the convergence of costs as in (5.37), completing the proof of the theorem.

6 Proofs of Results from Section 5

In this section we provide the proofs of Lemmas 5.1, 5.2 and Proposition 5.1

Proof of Lemma 5.1: We start by analyzing the evolution of the value function $V^h_\nu$. Clearly, $V^h_\nu(T, x) = g(\nu(T), x)$ for all $x \in S^h_0$. Using backwards induction, we get that for any $(t,x,\nu) \in \mathbb{T}^h \times S^h_0 \times \mathcal{P}_{T,L}$, one has from the definition of $\hat{\alpha}^{h,\nu}$ (cf. (3.9)), $q^h$ (see (5.11)) and (5.13)–(5.14),

$$V^h_\nu(t, x) = \min_{u \in \mathcal{U}} \left\{ \Delta^h f(t, \nu(t), x, u) + q^h(t, \nu(t), x, x + h) \right\}$$

$$= \Delta^h f(t, \nu(t), x, \alpha^{h,\nu}(t, x)) + q^h(t, \nu(t), \alpha^{h,\nu}(t, x); x, x + h),$$

where $\alpha^{h,\nu}$ is as in (5.14). The above identity in particular shows that $\hat{\alpha}^{h,\nu}$ gives an optimal feedback control.

We now use (6.1) to prove (5.15). Define the following $h$-th finite differences for $(t,x) \in \mathbb{T}^h \times S^h_0$,

$$\mathcal{D}^h_{t} V^h_\nu(t, x) \doteq \frac{1}{\Delta^h} (V^h_\nu(t + \Delta^h, x) - V^h_\nu(t, x)),$$

$$\mathcal{D}^h_{xx} V^h_\nu(t, x) \doteq \frac{1}{h^2} (V^h_\nu(t + \Delta^h, x + h) - 2V^h_\nu(t + \Delta^h, x) + V^h_\nu(t + \Delta^h, x - h)).$$

Simplifying (6.1) by using (5.12), we get that

$$\mathcal{D}^h_{t} V^h_\nu(t, x) = -H(t, \nu(t), x, \mathcal{D}^h_x V^h_\nu(t, x)) - \frac{1}{2} \sigma^2 \mathcal{D}^h_{xx} V^h_\nu(t, x).$$
Notice also that $X^{h,\nu}(t + \Delta h) \neq X^{h,\nu}(t)$ if and only if $R^{h,\nu}(t + \Delta h) - R^{h,\nu}(t) = Y^{h,\nu}(t + \Delta h) - Y^{h,\nu}(t) = 0$, in which case, one has

$$V^h(t + \Delta h, X^{h,\nu}(t + \Delta h)) - V^h(t, X^{h,\nu}(t))$$

$$= \mathcal{D}_t V^h(t, X^{h,\nu}(t)) \Delta h + \mathcal{D}_x V^h(t, X^{h,\nu}(t))(X^{h,\nu}(t + \Delta h) - X^{h,\nu}(t)) + \frac{1}{2} h^2 \mathcal{D}_{xx} V^h(t, X^{h,\nu}(t)).$$

This can be easily verified by considering separately the cases $X^{h,\nu}(t + \Delta h) = X^{h,\nu}(t) \pm h$. In case that $R^{h,\nu}(t + \Delta h) - R^{h,\nu}(t) = h$, that is $X^{h,\nu}(t + \Delta h) = X^{h,\nu}(t) = L$, (5.12) implies that

$$V^h(t + \Delta h, X^{h,\nu}(t + \Delta h)) - V^h(t, X^{h,\nu}(t))$$

$$= V^{h,\nu}(t + \Delta h, L) - V^{h,\nu}(t + \Delta h, h) + V^{h,\nu}(t + \Delta h, h + \Delta h) - V^{h,\nu}(t, L)$$

$$= -r(t + \Delta h)h + \mathcal{D}_t V^h(t, X^{h,\nu}(t)) \Delta h + \mathcal{D}_x V^h(t, X^{h,\nu}(t))h + \frac{1}{2} h^2 \mathcal{D}_{xx} V^h(t, X^{h,\nu}(t))$$

and finally, in case that $Y^{h,\nu}(t + \Delta h) - Y^{h,\nu}(t) = h$, one similarly has

$$V^h(t + \Delta h, X^{h,\nu}(t + \Delta h)) - V^h(t, X^{h,\nu}(t))$$

$$= -y(t + \Delta h)h + \mathcal{D}_t V^h(t, X^{h,\nu}(t)) \Delta h + \mathcal{D}_x V^h(t, X^{h,\nu}(t))(-h) + \frac{1}{2} h^2 \mathcal{D}_{xx} V^h(t, X^{h,\nu}(t))$$

From the last two equalities, we get that,

$$V^h(t + \Delta h, X^{h,\nu}(t + \Delta h)) - V^h(t, X^{h,\nu}(t))$$

$$= -H(t, \nu(t), X^{h,\nu}(t), \mathcal{D}_t V^h(t, X^{h,\nu}(t))(X^{h,\nu}(t + \Delta h) - X^{h,\nu}(t))$$

$$- (y(t + h) + \mathcal{D}_x V^h(t, X^{h,\nu}(t))(Y^{h,\nu}(t + \Delta h) - Y^{h,\nu}(t))$$

$$- (r(t + \Delta h) - \mathcal{D}_x V^h(t, X^{h,\nu}(t))(R^{h,\nu}(t + \Delta h) - R^{h,\nu}(t))$$

$$= -f(t, \nu(t), X^{h,\nu}(t), \alpha^{h,\nu}(t, X^{h,\nu}(t))) \Delta h + \sigma \mathcal{D}_x V^h(t, X^{h,\nu}(t))(B^{h,\nu}(t + \Delta h) - B^{h,\nu}(t))$$

$$- y(t + \Delta h)(Y^{h,\nu}(t + \Delta h) - Y^{h,\nu}(t)) - r(t + \Delta h)(R^{h,\nu}(t + \Delta h) - R^{h,\nu}(t)).$$

Summing up the terms over $t \in \mathbb{T}^h$, one gets (5.15). We will postpone the proof of (5.16) to the end of the paper. \hfill \Box

**Proof of Lemma 5.2** Fix $t \in [0, T]$, $\eta, \eta' \in \mathcal{P}([0, L])$, $x, x' \in [0, L]$, and $p, p' \in \mathbb{R}$. Denote $\alpha = \hat{\alpha}(t, \eta, x, p)$ and $\alpha' = \hat{\alpha}(t, \eta', x', p')$. Recall the definition of $h$ from (3.7). By (5.19), (5.20), and the definition of $\alpha'$, we get that

$$h(t, \eta', x', \alpha, p') \geq h(t, \eta', x', \alpha', p') \geq h(t, \eta', x', \alpha, p') + (\alpha' - \alpha)h_\alpha(t, \eta', x', \alpha, p') + c_m|\alpha' - \alpha|^2.$$

From the minimizing property of $\alpha$ we see that $(\alpha' - \alpha)h_\alpha(t, \eta, x, \alpha, p) \geq 0$. Subtracting this term from the right side of the above

$$c_m|\alpha' - \alpha|^2 \leq |\alpha' - \alpha| \cdot |h_\alpha(t, \eta', x', \alpha, p') - h_\alpha(t, \eta, x, \alpha, p)|$$

$$\leq c|\alpha' - \alpha||W_2(\eta, \eta') + |x - x'| + |p - p'||,$$

where $c = c_1 + \sup_{t \in [0, T]} |b_2(t)|$ and the second inequality follows by (5.19) and (5.21). Result follows on dividing both sides by $|\alpha' - \alpha|/c_m$. \hfill \Box
Proof of Proposition 5.1: We begin by introducing a coupling between two optimally controlled chains, one associated with $\nu$ and the other with $\nu'$.

**Coupling.** Fix $x \in [0, L]$ and $\nu, \nu' \in \mathcal{P}_{T,L}$. Let $\{X^\nu_n\}$ and $\{X'^\nu_n\}$ be the Markov chains from Construction 5.1 associated with the parameter $h$ and the optimal strategies given by (5.14). Denote by $\Sigma^\nu = (n^\nu, F^\nu, B^\nu, X^\nu, Y^\nu, R^\nu)$ and $\Sigma'^\nu = (n'^\nu, F'^\nu, B'^\nu, X'^\nu, Y'^\nu, R'^\nu)$ the processes that were defined immediately after Construction 5.1 where we suppressed the index $h$ since it is fixed in the rest of the proof. Also, denote

$$b'(t) = b(l^h(t), \nu(l^h(t)), X'^\nu(t), \hat{a}'(t, X'(t))),$$

where $l^h(t) = \lfloor t/\Delta^h \rfloor \Delta^h$. Similarly define $b''$. We now define a coupling of the chains through a time change of an underlying Markov chain $\{(Z''^\nu_n, Z''^{\nu'})\}$. The main idea in the construction of the latter Markov chain is to keep track of the proper time. Whenever an ‘instantaneous jump’ occurs for only one of the $Z$-processes, the other process has a degenerate step, that is, it remains at the same position. Therefore, we use two sequences of times. The first, which we refer as time instants, $\{t_n\}$ has the same role as in (5.4). The second is referred as time steps and denoted by $\{(N''^\nu_n, N''^{\nu'})\}$. Each of the components counts how many non-degenerate steps the respective $Z$ process has done so far.

Set $Z''^\nu_0 = Z''^{\nu'}_0 = x_0$, $t_0 = 0$, and $N''^\nu_0 = N''^{\nu'}_0 = 0$. Having defined for $i = 0, 1, \ldots, n$ time instants $t''_i < T$, time steps $N''^\nu_i, N''^{\nu'}_i \in \mathbb{N}$, and random variables $Z''_i, Z''^{\nu'}_i$ with values in $\mathbb{S}^h$, define them for the $(n+1)$-th step as follows.

- If $Z''^\nu_n, Z''^{\nu'}_n \notin \{-h, L+h\}$, then

$$\begin{aligned}
(Z''^\nu_{n+1}, Z''^{\nu'}_{n+1}) &= (Z''^\nu_n, Z''^{\nu'}_n) + \\
&= \begin{cases}
(h, h), & \text{w.p. } (h \min \{b''(t_n), b''(t'_n)\} + \sigma^2)/(2\sigma^2), \\
(h, -h), & \text{w.p. } \frac{h(b''(t_n) - b''(t'_n))^+/(2\sigma^2)}, \\
(-h, h), & \text{w.p. } \frac{h(b''(t_n) - b''(t'_n))^-//(2\sigma^2)}, \\
(-h, -h), & \text{w.p. } (-h \max \{b''(t_n), b''(t'_n)\} + \sigma^2)/(2\sigma^2)
\end{cases}
\end{aligned}$$

and

$$t_{n+1} = t_n + \Delta^h, \quad N''^\nu_{n+1} = N''^\nu_n + 1, \quad \text{and} \quad N''^{\nu'}_{n+1} = N''^{\nu'}_n + 1,$$

where $x^+ = \max\{0, x\}$ and $x^- = \max\{0, -x\}$.

- If $Z''^\nu_n \notin \{-h, L+h\}$ and $Z''^{\nu'}_n \in \{-h, L+h\}$ then

$$\begin{aligned}
(Z''^\nu_{n+1}, Z''^{\nu'}_{n+1}) &= (Z''^\nu_n, Z''^{\nu'}_n) + \\
&= \begin{cases}
(0, h(-1)^\frac{3}{2}(Z''^{\nu'}_n - L+h)), \\
(h, h), & \text{w.p. } (h \min \{b''(t_n), b''(t'_n)\} + \sigma^2)/(2\sigma^2), \\
(h, -h), & \text{w.p. } \frac{h(b''(t_n) - b''(t'_n))^+/(2\sigma^2)}, \\
(-h, h), & \text{w.p. } \frac{h(b''(t_n) - b''(t'_n))^-//(2\sigma^2)}, \\
(-h, -h), & \text{w.p. } (-h \max \{b''(t_n), b''(t'_n)\} + \sigma^2)/(2\sigma^2)
\end{cases}
\end{aligned}$$

and

$$t_{n+1} = t_n, \quad N''^\nu_{n+1} = N''^\nu_n, \quad \text{and} \quad N''^{\nu'}_{n+1} = N''^{\nu'}_n + 1.$$

The transition probabilities when $Z''^\nu_n \notin \{-h, L+h\}$ and $Z''^{\nu'}_n \in \{-h, L+h\}$ are defined similarly.
We also define the corresponding ‘unconstrained’ increment as

\[(Z_{n+1}^{\nu}, Z_{n+1}^{\nu'}) = (Z_n^{\nu}, Z_n^{\nu'}) + \left(h(-1)^{2(\nu_n-L+\nu)\bmod 2}, h(-1)^{1(\nu_1+\nu_n-L)}\right)\]

and

\[t_{n+1} = t_n, \quad N_{n+1}^{\nu} = N_n^{\nu} + 1, \quad \text{and} \quad N_{n+1}^{\nu'} = N_n^{\nu'} + 1.\]

For every \(n \in \mathbb{N}\), set

\[X_n^{\nu} = Z_{M_n}^{\nu} \quad \text{and} \quad X_n^{\nu'} = Z_{M_n'}^{\nu'},\]

where \(M_n^{\nu} = \max\{m : N_m^{\nu} \leq n\}\) and similarly \(M_n^{\nu'}\) is defined.

With the above construction \(\{X_n^{\nu}\}\) and \(\{X_n^{\nu'}\}\) are controlled Markov chains constructed using the optimal feedback controls \(\hat{a}^{\nu}\) and \(\hat{a}^{\nu'}\) respectively, given on the same probability space. Also relationships (5.5)–(5.8) are satisfied by \(\Sigma^{\nu}\) and \(\Sigma^{\nu'}\). The above coupling of the two processes gives the joint evolution \((X^{\nu}(t), X^{\nu'}(t))_{0 \leq t \leq T}\) as follows. \(X^{\nu}(0) = X^{\nu'}(0) = x_0\) and for every \(t \in \mathbb{T}^h\),

\[(X^{\nu}(t + \Delta^h), X^{\nu'}(t + \Delta^h)) - (X^{\nu}(t), X^{\nu'}(t)) = \begin{cases} (h1_{\{X^{\nu}(t) \neq L\}}, h1_{\{X^{\nu'}(t) \neq L\}}, & \text{w.p. } \frac{(h\min\{b^{\nu}(t), b^{\nu'}(t)\} + \sigma^2)}{2\sigma^2}, \\
(h1_{\{X^{\nu}(t) \neq L\}}, -h1_{\{X^{\nu'}(t) \neq 0\}}), & \text{w.p. } \frac{h(\Delta b(t))^+}{2\sigma^2}, \\
(-h1_{\{X^{\nu}(t) \neq 0\}}, h1_{\{X^{\nu'}(t) \neq L\}}), & \text{w.p. } \frac{h(\Delta b(t))^-}{2\sigma^2}, \\
(-h1_{\{X^{\nu}(t) \neq 0\}}, -h1_{\{X^{\nu'}(t) \neq 0\}}), & \text{w.p. } \frac{-h \max\{b^{\nu}(t), b^{\nu'}(t)\} + \sigma^2}{2\sigma^2}, \end{cases}\]

where

\[\Delta b(t) = b^{\nu}(t) - b^{\nu'}(t), \quad t \in [0, T].\]

We also define the corresponding ‘unconstrained’ increment as

\[(Z^{\nu}(t + \Delta^h), Z^{\nu'}(t + \Delta^h)) - (X^{\nu}(t), X^{\nu'}(t)) = \begin{cases} (h, h), & \text{w.p. } \frac{(h\min\{b^{\nu}(t), b^{\nu'}(t)\} + \sigma^2)}{2\sigma^2}, \\
(h, -h), & \text{w.p. } \frac{h(\Delta b(t))^+}{2\sigma^2}, \\
(-h, h), & \text{w.p. } \frac{h(\Delta b(t))^-}{2\sigma^2}, \\
(-h, -h), & \text{w.p. } \frac{-h \max\{b^{\nu}(t), b^{\nu'}(t)\} + \sigma^2}{2\sigma^2}, \end{cases}\] (6.2)

Bounding \(W_1^2(\Phi^{h}(\nu), \Phi^{h}(\nu'))\). Denote \(\Delta X(t) = X^{\nu}(t) - X^{\nu'}(t)\). The processes \(\Delta B, \Delta R\) and \(\Delta Y\) are defined similarly. Note that

\[W_1^2(\Phi^{h}(\nu), \Phi^{h}(\nu')) \leq \mathbb{E}[|\Delta X|^2] .\]

We now estimate \(\mathbb{E}[|\Delta X|^2]\). Recall that \(\Delta X(0) = 0\). From (5.5) and Lemma 2.1

\[|\Delta X|_T \leq |\Delta X|_T + |\Delta Y|_T + |\Delta R|_T \leq c \left( \int_0^T |\Delta b(s)| ds + \sigma |\Delta B|_T \right).\]
Therefore,
\[
\mathbb{E}[|\Delta X|^2_{T}] \leq 2c_{2}^{2}\mathbb{E}\left[\left(\int_{0}^{T}|\Delta b(s)|ds\right)^{2} + \sigma^{2}|\Delta B|_{T}^{2}\right].
\] (6.3)

We now estimate the second term on the right side. By using the martingale property of \(B^{h,\nu}(t) - B^{h,\nu'}(t)\) and Doob’s inequality,
\[
\mathbb{E}[|\Delta B|_{T}^{2}] \leq 4\mathbb{E}\left[\sum_{s \in \mathbb{T}^{h}}(\Delta B(s + \Delta) - \Delta B(s))^{2}\right].
\] (6.4)

From (5.7) and (5.9),
\[
\sigma|\Delta B(s + \Delta) - \Delta B(s)| \leq |(\Delta X + \Delta R - \Delta Y)(s + \Delta) - (\Delta X + \Delta R - \Delta Y)(s)| + \Delta^{|\Delta b(s)|}.
\]

If \((Z^{\nu}(t + \Delta) - X^{\nu}(t)) (Z^{\nu}(t + \Delta) - X^{\nu}(t)) > 0\), i.e. the unconstrained increments are of the same sign, then
\[
\left|(\Delta X + \Delta R - \Delta Y)(s + \Delta) - (\Delta X + \Delta R - \Delta Y)(s)\right|
= \left|(X^{\nu} + R^{\nu} - Y^{\nu})(s + \Delta) - (X^{\nu} + R^{\nu} - Y^{\nu})(s)\right|
- \left|(X^{\nu} + R^{\nu} - Y^{\nu})(s + \Delta) - (X^{\nu} + R^{\nu} - Y^{\nu})(s)\right| = 0.
\]

If the signs are different, i.e. \((Z^{\nu}(t + \Delta) - X^{\nu}(t)) (Z^{\nu}(t + \Delta) - X^{\nu}(t)) < 0\), then
\[
\left|(\Delta X + \Delta R - \Delta Y)(s + \Delta) - (\Delta X + \Delta R - \Delta Y)(s)\right| \leq 2h.
\]

Hence,
\[
\sigma|\Delta B(s + \Delta) - \Delta B(s)| \leq 2h\mathbb{1}_{E_{s}} + |\Delta b(s)|\Delta^{|h},
\] (6.5)

where
\[
\mathbb{E}\left[\left(\sigma(\Delta B(s + \Delta) - \Delta B(s))\right)^{2} \mid \mathcal{H}_{s}^{h}\right] \leq \mathbb{E}\left[\left(2h\mathbb{1}_{E_{s}} + |\Delta b(s)|\Delta^{|h}\right)^{2} \mid \mathcal{H}_{s}^{h}\right]
= 2h|\Delta b(s)|(\Delta^{|h} + C\Delta^{|h}),
\] (6.6)

where in the above expression, and in the rest of the proof, \(C\) refers to a finite positive constant that is independent of \(h\) and \(s, \nu, \nu'\) and which can change from one line to the next. Applying the above bound to (6.4), we get that for sufficiently small \(h\),
\[
\sigma^{2}\mathbb{E}[|\Delta B|_{T}^{2}] \leq 12\Delta^{h} \sum_{s \in \mathbb{T}^{h}} h\mathbb{E}[|\Delta b(s)|] = 12\mathbb{E}\left[\int_{0}^{T} h|\Delta b(s)|ds\right].
\] (6.7)
Combining this with (6.3) and using the inequality,

\[ h \int_{0}^{T} |\Delta b(s)| ds \leq \frac{1}{2} \left[ h^2 T^{1/2} + T^{-1/2} \left( \int_{0}^{T} |\Delta b(s)| ds \right)^2 \right] \leq \frac{1}{2} T^{1/2} \left[ h^2 + \int_{0}^{T} |\Delta b(s)|^2 ds \right] \]  \hspace{1cm} (6.8)

we get that

\[ \mathbb{E}[|\Delta X|_{T}^2] \leq C \mathbb{E} \left[ \left( \int_{0}^{T} |\Delta b(s)| ds \right)^2 + h \int_{0}^{T} |\Delta b(s)| ds \right] \]

\[ \leq C T^{1/2} h^2 + C(T^{1/2} + T) \mathbb{E} \left[ \int_{0}^{T} (\Delta W(s))^2 ds + \int_{0}^{T} (\Delta X(s))^2 ds + \int_{0}^{T} (\Delta D^h V^h(s))^2 ds \right] \],

where for \( t \in [0, T] \)

\[ \Delta D^h V^h(t) = D^h V^h(t^h(t), X^\nu(t)) - D^h V^h(t^h(t), X^\nu(t)), \Delta W(t) = W_1(\nu(t^h(t)), \nu'(t^h(t))) \]

and the above inequality also uses the Lipschitz property of \( b \) (Assumption 3.1), the Lipschitz property of \( \hat{\alpha} \) (Assumption 5.1) and (5.14).

We now consider the last term on the right side of (6.9). We will show that for some \( \hat{T} \) that does not depend on \( h, \nu, \nu' \) and all \( T \leq \hat{T} \),

\[ \mathbb{E} \left[ \int_{0}^{T} (\Delta D^h V^h(s))^2 ds \right] \leq C \left( h^2 + \sup_{0 \leq s \leq T} (\Delta W(s))^2 + \mathbb{E}[|\Delta X|_{T}^2] \right) . \]  \hspace{1cm} (6.10)

The proof of this inequality follows along similar lines as the proof of the continuity of the function \( \Phi \) in the proof of Theorem 3.1 (see e.g. the proof of (3.27)). However, here we have a sequence of maps \( \Phi^h \) indexed by \( h \) and we seek a bound that is uniform in \( h \). Also, unlike in the proof of Theorem 3.1, here we do not have a convenient reference measure for estimates.

Define,

\[ \Delta V^h(s) = V^h_s(s, X^\nu(s)) - V^h_s(s, X^\nu(s)), \Delta g(T) = \Delta V^h(T) = g(\nu(T), X^\nu(T)) - g(\nu'(T), X^\nu'(T)), \]

\[ \Delta f(s) = f(t^h(s), \nu(t^h(s)), X^\nu(s)), \alpha^h(\nu(s), X^\nu(s)) = f(t^h(s), \nu'(s), X^\nu'(t^h(s)), \alpha^h(\nu(s), X^\nu(s)) \]

From (5.15) we get that

\[ \Delta V^h(0) + \sigma \sum_{s \in T^h} \Delta D^h V^h(s)(B^\nu(s + \Delta h) - B^\nu(s)) \\
= -\sigma \sum_{s \in T^h} D^h V^h_s(s, X^\nu(s))(\Delta B(s + \Delta h) - \Delta B(s)) + \Delta g(T) + \int_{0}^{T} \Delta f(s) ds \\
+ \int_{0}^{T} y(s)d(\Delta Y(s)) + \int_{0}^{T} r(s)d(\Delta R(s)). \]
By squaring both sides and taking expectations, we have
\[
(\Delta V^h(0))^2 + \mathbb{E}\left[\left(\sigma \sum_{s \in \mathbb{T}^h} \Delta D_x^h V^h(s)(B^\nu(s + \Delta^h) - B^\nu(s))\right)^2\right]
\]
(6.11)
\[
\leq 5 \left\{ \mathbb{E}\left[\left(\sigma \sum_{s \in \mathbb{T}^h} D_x^h V^h(s, X^\nu(s))(\Delta B(s + \Delta^h) - \Delta B(s))\right)^2\right] + \mathbb{E}[(\Delta g(T))^2] \right. \\
+ \mathbb{E}\left[\left(\int_0^T \Delta f(s)ds\right)^2\right] + \mathbb{E}\left[\left(\int_0^T g(s)d(\Delta Y(s))\right)^2\right] + \mathbb{E}\left[\left(\int_0^T r(s)d(\Delta R(s))\right)^2\right] \right\}.
\]

We now bound the terms in the inequality above. For any \(s \in \mathbb{T}^h\), we get by \eqref{6.13} and \eqref{5.7} that
\[
\mathbb{E}\left[\left(\Delta D_x^h V^h(s)(B^\nu(s + \Delta^h) - B^\nu(s))\right)^2 | \mathcal{H}_s\right] = (\Delta D_x^h V^h(s))^2(\Delta^h + C(\Delta^h)^2).
\]
Using once more the martingale property of \(B^\nu\) we get that, for sufficiently small \(h\),
\[
\mathbb{E}\left[\left(\sum_{s \in \mathbb{T}^h} \Delta D_x^h V^h(s)(B^\nu(s + \Delta^h) - B^\nu(s))\right)^2\right] = \sum_{s \in \mathbb{T}^h} \mathbb{E}\left[\left(\Delta D_x^h V^h(s)(B^\nu(s + \Delta^h) - B^\nu(s))\right)^2\right] \\
\geq \frac{1}{2} \mathbb{E}\left[\int_0^T (\Delta D_x^h V^h(s))^2ds\right].
\]
(6.12)

Recall that from \eqref{6.16}, \(|D_x^h V^h(\cdot, X^\nu(\cdot))|_T \leq c_d(M)\) whenever \(T \leq M\). Henceforth we will only consider \(T \in [0, M]\). Using \eqref{6.6}, we get that for sufficiently small \(h\),
\[
\mathbb{E}\left[\left(\sigma \sum_{s \in \mathbb{T}^h} D_x^h V^h(s, X^\nu(s))(\Delta B(s + \Delta^h) - \Delta B(s))\right)^2\right]
\]
(6.13)
\[
= \sum_{s \in \mathbb{T}^h} \mathbb{E}\left[\left(\sigma D_x^h V^h(s, X^\nu(s))(\Delta B(s + \Delta^h) - \Delta B(s))\right)^2\right] \\
\leq 3(c_d(M))^2\mathbb{E}\left[\int_0^T h|\Delta b(s)|ds\right] \\
\leq CT^{1/2}\mathbb{E}\left[h^2 + \int_0^T (\Delta W(s))^2ds + \int_0^T (\Delta X(s))^2ds + \int_0^T (\Delta D_x^h V^h(s))^2ds\right],
\]
where the last inequality follows by a similar inequality as in \eqref{6.8}.

From \eqref{3.8}, it is easy to see that
\[
\mathbb{E}[(\Delta g(T))^2] \leq C \left( (\Delta W(T))^2 + \mathbb{E}[(\Delta X(T))^2] \right),
\]
(6.14)
and that
\[
\mathbb{E}\left[\left(\int_0^T \Delta f(s)ds\right)^2\right] \leq C\mathbb{E}\left\{\left(\int_0^T \Delta W(s)ds\right)^2 + \left(\int_0^T \Delta X(s)ds\right)^2 + \left(\int_0^T \Delta D_x^h V^h(s)ds\right)^2\right\} \\
\leq C\mathbb{E}\left[\int_0^T (\Delta W(s))^2ds + \int_0^T (\Delta X(s))^2ds + \int_0^T (\Delta D_x^h V^h(s))^2ds\right].
\]
(6.15)
Using integration by parts and the boundedness of \( y \)

\[
\mathbb{E} \left[ \left( \int_0^T y(s)d(\Delta Y(s)) \right)^2 \right] 
\leq C \mathbb{E} \left[ \Delta Y_T^2 \right] 
\leq C \mathbb{E} \left[ \left| F_\nu + \sigma B_\nu - F_\nu' - \sigma B_\nu' \right|^2 \right] 
\leq C \left( \int_0^T |\Delta b(s)| ds \right)^2 + \sigma^2 |\Delta B_T^2| 
\leq CT^{1/2}h^2 + C(T^{1/2} + T) \mathbb{E} \left[ \int_0^T (\Delta W(s))^2 ds + \int_0^T (\Delta X(s))^2 ds + \int_0^T (\Delta D^b h V^h(s))^2 ds \right],
\]

where the inequality on the third line is from Lemma 2.1, the fourth line is from (5.9) and the last line uses (6.7) and (6.9). A similar bound holds for \( \mathbb{E} \left[ \left( \int_0^T r(s)d(\Delta R(s)) \right)^2 \right] \).

From (6.11)–(6.16), we have

\[
(\Delta V^h(0))^2 + \frac{1}{2} \sigma^2 \mathbb{E} \left[ \int_0^T (\Delta D^b h V^h(s))^2 ds \right] 
\leq C \left( (\Delta W(T))^2 + \mathbb{E}[(\Delta X(T))^2] \right) + CT^{1/2}h^2 + C(T^{1/2} + T) \mathbb{E} \left[ \int_0^T (\Delta W(s))^2 ds + \int_0^T (\Delta X(s))^2 ds + \int_0^T (\Delta D^b h V^h(s))^2 ds \right].
\]

Thus we can find a \( \hat{T}_1 \in (0, M) \) and \( \hat{h}_1 > 0 \) such that for all \( T \leq \hat{T}_1 \) and \( h \leq \hat{h}_1 \), (6.10) is satisfied. Together with (6.9) we now get that there exist \( \hat{T} \in (0, M) \), \( \hat{h} > 0 \) and \( q \in (0, 1) \) such that for every \( T \leq \hat{T} \) and \( h \in (0, \hat{h} \wedge \hat{T}) \),

\[
W_T^2(\Phi^h(\nu), \Phi^h(\nu')) \leq \mathbb{E}[|\Delta X_T^2|] \leq q \left( h^2 + \sup_{0 \leq s \leq T} (\Delta W(s))^2 \right) \leq q \left( h^2 + W_T^2(\nu, \nu') \right).
\]

\( \square \)

**Proof of (5.16):** Fix \( \nu \in \mathcal{P}_{T,L} \), \( x \neq x' \) in \( S_0^h \), and \( t_0 \in T^h \), which will be regarded as the initial time. As in the proof of Proposition 5.1, one can define a coupling of two processes on the same \( h \)-grid, both of which are driven by the same \( \nu \). The first one is denoted as \((X(s), Y(s), R(s), B(s), \alpha(s))_{t_0 \leq s \leq T}\), where its components are defined in \( 5.31 - 5.7 \) with \( X(t_0) = x \) and \( \alpha \) is the optimal policy for this process. The second process, denoted as \((X'(s), Y'(s), R'(s), B'(s), \alpha(s))_{t_0 \leq s \leq T}\) is also given by \( 5.31 - 5.7 \) using the same control process \( \{\alpha(s)\} \) as for the first one, except that the second one starts at \( x' \), i.e. \( X'(t_0) = x' \).

For every \( s \in [t_0, T] \), let \( \Delta X(s) = X(s) - X'(s) \) and

\[
\Delta b(s) = b(l^h(s), \nu(l^h(s)), X(s), \alpha(s)) - b(l^h(s), \nu(l^h(s)), X'(s), \alpha(s)).
\]

Processes \( \Delta Y(s) \) and \( \Delta R(s) \) are defined in a similar manner. By definition \( \Delta X(0) = x - x' \). The arguments that lead to (6.9) can also be applied here, but in fact they are simpler here.
since $\Delta W = 0$ and the controls are the same for both processes. Specifically,

$$
\mathbb{E} \left[ \sup_{t_0 \leq s \leq T} (|\Delta X(s)| + |\Delta R(s)| + |\Delta Y(s)|)^2 \right] \leq C(x - x')^2 + C_1(T - t_0)^{1/2} h^2 + C((T - t_0)^{1/2} + (T - t_0)) \mathbb{E} \left[ \int_{t_0}^T (\Delta X(s))^2 ds \right] \leq C(x - x')^2(1 + T^{1/2}) + C(T^{1/2} + T) \mathbb{E} \left[ \int_{t_0}^T (\Delta X(s))^2 ds \right].
$$

The second inequality is a consequence of the fact that since $x \neq x'$, $h \leq |x - x'|$. Therefore,

$$
\mathbb{E} \left[ \sup_{t_0 \leq s \leq T} (\Delta X(s))^2 \right] \leq C(x - x')^2(1 + T^{1/2}) + C(T^{1/2} + T) \mathbb{E} \left[ \int_{t_0}^T (\Delta X(s))^2 ds \right].
$$

By Grönwall’s inequality, \( \mathbb{E} \left[ \sup_{t_0 \leq s \leq T} (\Delta X(s))^2 \right] \leq c_T |x - x'|^2, \) where

$$
c_T = C(1 + T^{1/2}) \exp \left\{ C_1(T^{3/2} + T^2) \right\}.
$$

Using the above bound we have that the left side in (6.18) can be bounded above by

$$
C(x - x')^2(1 + T^{1/2}) + C(T^{1/2} + T) \mathbb{E} \left[ \sup_{t_0 \leq s \leq T} (\Delta X(s))^2 \right] \leq \tilde{c}_T |x - x'|^2,
$$

where $\tilde{c}_T := (4CL^2(1 + T^{1/2}) + C(T^{1/2} + T)CT_T)^{1/2}$, and so,

$$
\mathbb{E} \left[ \sup_{t_0 \leq s \leq T} (|\Delta X(s)| + |\Delta Y(s)| + |\Delta R(s)|) \right] \leq (\tilde{c}_T)^{1/2}|x - x'|,
$$

Consequently, using integration by parts as in (6.16) and Lipschitz property of $f$ and $g$, we get that

$$
V_\nu^h(t, x') - V_\nu^h(t, x) \leq J^{h, \nu}(t, x', \alpha) - J^{h, \nu}(t, x, \alpha)
$$

$$
\leq \mathbb{E} \left[ |g(\nu(T), X'(T)) - g(\nu(T), X(T))| + \left| \int_{t_0}^T y(s) d(\Delta Y(s)) \right| + \left| \int_{t_0}^T r(s) d(\Delta Y(s)) \right| + \int_{t_0}^T |f(l^h(s), \nu(l^h(s)), X'(s), \alpha(s)) - f(l^h(s), \nu(l^h(s)), X(s), \alpha(s))| ds \right]
$$

$$
\leq \tilde{c}_T |x - x'|,
$$

where $\tilde{c}_T$ depends on the parameters $c_L, c_S$, the bounds on $y$ and $r$, and the terminal time $T$. By reversing the roles of the processes we get that

$$
|V_\nu^h(t, x) - V_\nu^h(t, x')| \leq \tilde{c}_T |x - x'|
$$

and the result follows.

\begin{proof}

\end{proof}

\textbf{Acknowledgement:} We thank the anonymous referees and the AE for insightful comments, which helped us improve our paper.

45
References

[1] Y. Achdou and I. Capuzzo-Dolcetta. Mean field games: Numerical methods. *SIAM J. Numer. Anal.*, 48(3):1136–1162, 2010.

[2] C. D. Aliprantis and K. C. Border. *Infinite dimensional analysis: A hitchhiker’s guide*. Springer, Berlin, third edition, 2006.

[3] N. Antunes, C. Fricker, P. Robert, and D. Tibi. Stochastic networks with multiple stable points. *The Annals of Probability*, 36(1):255–278, 2008.

[4] B. Ata, J. M. Harrison, and L. A. Shepp. Drift rate control of a Brownian processing system. *Ann. Appl. Probab.*, 15(2):1145–1160, 2005.

[5] F. Baccelli, F. Karpelevich, M. Y. Kelbert, A. Puhalskii, A. Rybko, and Y. M. Suhov. A mean-field limit for a class of queueing networks. *Journal of Statistical Physics*, 66(3-4):803–825, 1992.

[6] E. Bayraktar, U. Horst, and R. Sircar. A limit theorem for financial markets with inert investors. *Math. Oper. Res.*, 31(4):789–810, 2006.

[7] E. Bayraktar, U. Horst, and R. Sircar. Queuing theoretic approaches to financial price fluctuations. *Handbooks in Operations Research and Management Science*, 15:637–677, 2007.

[8] E. Bayraktar and M. Ludkovski. Optimal trade execution in illiquid markets. *Math. Finance*, 21(4):681–701, 2011.

[9] E. Bayraktar and M. Ludkovski. Liquidation in limit order books with controlled intensity. *Math. Finance*, 24(4):627–650, 2014.

[10] E. Bayraktar and Y. Zhang. A rank-based mean field game in the strong formulation. *Electron. Commun. Probab.*, 21:Paper No. 72, 12, 2016.

[11] M. Benaim and J.-Y. Le Boudec. A class of mean field interaction models for computer and communication systems. *Performance Evaluation*, 65(11):823–838, 2008.

[12] P. Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[13] J. Blair, P. Johnson, and P. Duck. Analysis of optimal liquidation in limit order books. http://eprints.ma.man.ac.uk/2299/01/covered/MIMS_ep2014_23.pdf, 2015. Preprint.

[14] A. Bobbio, M. Gribaudo, and M. Telek. Analysis of large scale interacting systems by mean field method. In *Quantitative Evaluation of Systems, 2008. QEST’08. Fifth International Conference on*, pages 215–224. IEEE, 2008.

[15] V. S. Borkar. Optimal control of diffusion processes. volume 203 of *Pitman Research Notes in Mathematics Series*, pages vi+196. Longman Scientific & Technical, Harlow; copublished in the United States with John Wiley & Sons, Inc., New York, 1989.

[16] K. A. Borovkov. Propagation of chaos for queueing networks. *Theory Probab. Appl.*, 42(3):385–394, 1998.

[17] M. Bowe, S. Hyde, and I. Johnson. Determining the intensity of buy and sell limit order submissions: A look at the market preopening period.

[18] A. Budhiraja and E. Friedlander. Diffusion approximations for controlled weakly interacting large finite state systems with simultaneous jumps. *preprint, arXiv:1603.09001*, 2016.

[19] A. Budhiraja, A. P. Ghosh, and C. Lee. Ergodic rate control problem for single class queueing networks. *SIAM J. Control Optim.*, 49(4):1570–1606, 2011.
[20] P. Cardaliaguet. Notes on mean field games. 
https://www.ceremade.dauphine.fr/ cardalia/MFG20130420.pdf/, 2013.
[21] R. Carmona and F. Delarue. Probabilistic analysis of mean-field games. SIAM J. Control Optim., 51(4):2705–2734, 2013.
[22] R. Carmona and D. Lacker. A probabilistic weak formulation of mean field games and applications. Ann. Appl. Probab., 25(3):1189–1231, 2015.
[23] H. Chen and D. D. Yao. Fundamentals of Queuing Networks: Performance, Asymptotics, and Optimization. Springer: Berlin, 2001.
[24] J. G. Dai and R. J. Williams. Existence and uniqueness of semimartingale reflecting brownian motions in convex polyhedrons. Theory Probab. Appl., 40(1):1–40, 1996.
[25] N. El Karoui, D. Hû Nguyen, and M. Jeanblanc-Picqué. Compactification methods in the control of degenerate diffusions: Existence of an optimal control. Stochastics, 20(3):169–219, 1987.
[26] K. W. Fendick and M. A. Rodrigues. Asymptotic analysis of adaptive rate control for diverse sources with delayed feedback. Information Theory, IEEE Transactions on, 40(6):2008–2025, 1994.
[27] M. Fischer. On the connection between symmetric n-player games and mean field games. Ann. Appl. Probab., 2016. to appear.
[28] W. H. Fleming. Generalized solutions in optimal stochastic control. Technical report, DTIC Document, 1976.
[29] H. Föllmer and U. Horst. Convergence of locally and globally interacting Markov chains. Stochastic Process. Appl., 96(1):99–121, 2001.
[30] R. Gibbens, P. Hunt, and F. Kelly. Bistability in communication networks. Disorder in physical systems, pages 113–128, 1990.
[31] D. A. Gomes, J. Mohr, and R. R. Souza. Continuous time finite state mean field games. Appl. Math. Optim., 68(1):99–143, 2013.
[32] R. Gopalakrishnan, S. Doroudi, A. R. Ward, and A. Wierman. Routing and staffing when servers are strategic. Oper. Res., 2016. to appear.
[33] C. Graham. Chaoticity on path space for a queueing network with selection of the shortest queue among several. Journal of Applied Probability, 37(1):198–211, 2000.
[34] C. Graham and P. Robert. Interacting multi-class transmissions in large stochastic networks. The Annals of Applied Probability, 19(6):2334–2361, 2009.
[35] O. Guéant, J.-M. Lasry, and P.-L. Lions. Mean field games and applications. In Paris-Princeton Lectures on Mathematical Finance 2010, volume 2003 of Lecture Notes in Math., pages 205–266. Springer, Berlin, 2011.
[36] O. Guéant and C.-A. Lehalle. General intensity shapes in optimal liquidation. Math. Finance, 25(3):457–495, 2015.
[37] Y. Hu and S. Peng. A stability theorem of backward stochastic differential equations and its application. C. R. Acad. Sci. Paris Sér. I Math., 324(9):1059–1064, 1997.
[38] M. Huang, P. E. Caines, and R. P. Malhamé. The Nash certainty equivalence principle and McKean-Vlasov systems: An invariance principle and entry adaptation. In Decision and Control, 2007 46th IEEE Conference on, pages 121–126. IEEE, 2007.
[39] M. Huang, R. P. Malhamé, and P. E. Caines. Large population stochastic dynamic games: Closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. Commun. Inf. Syst., 6(3):221–251, 2006.

[40] P. J. Hunt and T. G. Kurtz. Large loss networks. Stochastic Process. Appl., 53(2):363–378, 1994.

[41] A. Joffe and M. Métivier. Weak convergence of sequences of semimartingales with applications to multitype branching processes. Adv. in Appl. Probab, pages 20–65, 1986.

[42] P. M. Kotelenez and T. G. Kurtz. Macroscopic limits for stochastic partial differential equations of McKean-Vlasov type. Probab. Theory Related Fields, 146(1-2):189–222, 2010.

[43] L. Kruk, J. Lehoczky, K. Ramanan, and S. Shreve. An explicit formula for the Skorokhod map on [0, a]. Ann. Probab., 35(5):1740–1768, 2007.

[44] H. J. Kushner. Heavy traffic analysis of controlled queueing and communication networks, volume 47 of Applications of Mathematics (New York). Springer-Verlag, New York, 2001. Stochastic Modelling and Applied Probability.

[45] H. J. Kushner and P. G. Dupuis. Numerical methods for stochastic control problems in continuous time, volume 24 of Applications of Mathematics (New York). Springer-Verlag, New York, 1992.

[46] A. Lachapelle, J. Lasry, C. Lehalle, and P. Lions. Efficiency of the price formation process in presence of high frequency participants: a mean field game analysis. http://arxiv.org/abs/1305.6323, 2015.

[47] D. Lacker. A general characterization of the mean field limit for stochastic differential games. Probab. Theory Related Fields, pages 1–68, 2015.

[48] D. Lacker. Mean field games via controlled martingale problems: existence of Markovian equilibria. Stochastic Process. Appl., 125(7):2856–2894, 2015.

[49] J.-M. Lasry and P.-L. Lions. Jeux `a champ moyen. I. Le cas stationnaire. C. R. Math. Acad. Sci. Paris, 343(9):619–625, 2006.

[50] J.-M. Lasry and P.-L. Lions. Jeux `a champ moyen. II. Horizon fini et contrôle optimal. C. R. Math. Acad. Sci. Paris, 343(10):679–684, 2006.

[51] J.-M. Lasry and P.-L. Lions. Mean field games. Jpn. J. Math., 2(1):229–260, 2007.

[52] J. Li, R. Bhattacharyya, S. Paul, S. Shakkottai, and V. Subramanian. Incentivizing sharing in realtime d2d streaming networks: A mean field game perspective. In Computer Communications (INFOCOM), 2015 IEEE Conference on, pages 2119–2127. IEEE, 2015.

[53] G. M. Lieberman. Second order parabolic differential equations. World Scientific Publishing Co., Inc., River Edge, NJ, 1996.

[54] M. Manjrekar, V. Ramaswamy, and S. Shakkottai. A mean field game approach to scheduling in cellular systems. In INFOCOM, 2014 Proceedings IEEE, pages 1554–1562. IEEE, 2014.

[55] J. P. Schauder. Der fixpunktsatz in funktionalthâumen. Studia Math, 2:171–180, 1930.

[56] D. R. Smart. Fixed point theorems. Cambridge University Press, London-New York, 1974. Cambridge Tracts in Mathematics, No. 66.

[57] A.-S. Sznitman. Topics in propagation of chaos. In Ecole d’été de probabilités de Saint-Flour XIX1989, pages 165–251. Springer, 1991.

[58] A. B. Tsybakov. Introduction to nonparametric estimation. Springer Series in Statistics. Springer, New York, 2009. Revised and extended from the 2004 French original, Translated by Vladimir Zaiats.
[59] C. Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.

[60] N. D. Vvedenskaya and Y. M. Suhov. Dobrushin’s Mean-Field Approximation for a Queue With Dynamic Routing. Research Report RR-3328, INRIA, 1997. Projet MEVAL.

[61] P. Wiecek, E. Altman, and A. Ghosh. Mean-field game approach to admission control of an $M/M/\infty$ queue with shared service cost. *Dyn. Games Appl.*, pages 1–29, 2015.

Erhan Bayraktar  
Department of Mathematics  
University of Michigan  
Ann Arbor, MI 48109, USA  
email: erhan@umich.edu  
web: www-personal.umich.edu/~erhan/

Amarjit Budhiraja  
Department of Statistics and Operations Research  
University of North Carolina  
Chapel Hill, NC 27599, USA  
email: budhiraj@email.unc.edu web: [HTTP://WWW.UNC.EDU/~ BUDHIRAJ/](http://www.unc.edu/~budhiraj/)

Asaf Cohen  
Department of Mathematics  
University of Michigan  
Ann Arbor, MI 48109, USA  
email: shloshim@gmail.com  
web: [HTTPS://SITES.GOOGLE.COM/SITE/ASAFCOHERNTA/](https://sites.google.com/site/asafcohenau/)