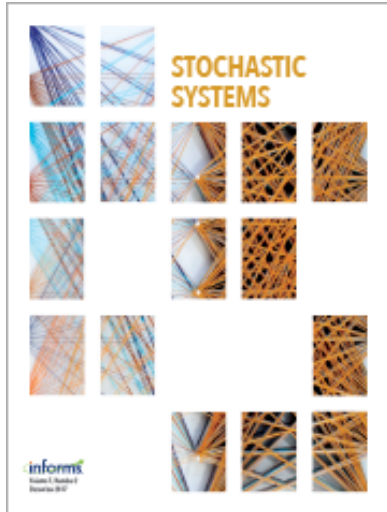




## Stochastic Systems

Publication details, including instructions for authors and subscription information:  
<http://pubsonline.informs.org>

### An Ergodic Control Problem for Many-Server Multiclass Queueing Systems with Cross-Trained Servers

Anup Biswas

To cite this article:

Anup Biswas (2017) An Ergodic Control Problem for Many-Server Multiclass Queueing Systems with Cross-Trained Servers. *Stochastic Systems* 7(2):264–288. <https://doi.org/10.1287/stsy.2017.0002>

This work is licensed under a Creative Commons Attribution 4.0 International License. You are free to copy, distribute, transmit and adapt this work, but you must attribute this work as “*Stochastic Systems*. Copyright © 2017 The Author(s). <https://doi.org/10.1287/stsy.2017.0002>, used under a Creative Commons Attribution License: <https://creativecommons.org/licenses/by/4.0/>.”

Copyright © 2017, The Author(s)

Please scroll down for article—it is on subsequent pages



With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes. For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

# An Ergodic Control Problem for Many-Server Multiclass Queueing Systems with Cross-Trained Servers

Anup Biswas<sup>a</sup>

<sup>a</sup>Department of Mathematics, Indian Institute of Science Education and Research, Pune 411008, India

Contact: [anup@iiserpune.ac.in](mailto:anup@iiserpune.ac.in) (AB)

Received: December 2015

Accepted: July 2017

Published Online in Articles in Advance:  
December 6, 2017

MSC2010 Subject Classification: Primary  
93E20; Secondary 60H30, 35J60.

<https://doi.org/10.1287/stsy.2017.0002>

Copyright: © 2017 The Author(s)

**Abstract.** A Markovian queueing network is considered with  $d$  independent customer classes and  $d$  server pools in Halfin-Whitt regime. Class  $i$  customers has priority for service in pool  $i$  for  $i = 1, \dots, d$ , and may access some other pool if the pool has an idle server and all the servers in pool  $i$  are busy. We formulate an ergodic control problem where the running cost is given by a non-negative convex function with polynomial growth. We show that the limiting controlled diffusion is modelled by an action space which depends on the state variable. We provide a complete analysis for the limiting ergodic control problem and establish asymptotic convergence of the value functions for the queueing model.

**History:** Former designation of this paper was SSY-2015-209.

**Open Access Statement:** This work is licensed under a Creative Commons Attribution 4.0 International License. You are free to copy, distribute, transmit and adapt this work, but you must attribute this work as “Stochastic Systems. Copyright © 2017 The Author(s). <https://doi.org/10.1287/stsy.2017.0002>, used under a Creative Commons Attribution License: <https://creativecommons.org/licenses/by/4.0/>.”

**Funding:** Supported in part by an INSPIRE faculty fellowship.

**Keywords:** multi-class Markovian queues • reneging/abandonment • Halfin-Whitt (QED) regime • heavy-traffic • long time-average control • scheduling control • stable Markov optimal control • Hamilton-Jacobi-Bellman equation • asymptotic optimality

## 1. Introduction

In this article we consider a queueing system consisting of  $d$  independent customer classes and  $d$  server pools (or stations). Each server pool contains  $n$  identical servers. Customers of class  $i$  have priority for service in pool  $i$  and this priority is of preemptive type. Therefore a newly arrived job of class  $i$  at time  $t$  would preempt the service of a class  $j$  job,  $j \neq i$ , if there is a class  $j$  job receiving service from pool  $i$  at time  $t$ . A customer from class  $j$  may access service from pool  $i$ ,  $j \neq i$ , if and only if there is an empty server in the pool  $i$  and all the servers in the pool  $j$  are busy. Therefore service stations are *cross-trained* to serve nonpriority customers when its own priority customer class is *underloaded*. Customers are also allowed to renege the system from the queue. The arrival of customers are given by  $d$ -independent Poisson processes with parameter  $\lambda_i^n$ ,  $i = 1, \dots, d$ . By  $\mu_{ij}^n$  we denote the service rate of class  $i$  customers at station  $j$ . We shall use  $\mu_i$  instead of  $\mu_{ii}$  for  $i = 1, \dots, d$ . The network is assumed to work under Halfin-Whitt setting in the sense that

$$\lim_{n \rightarrow \infty} \sqrt{n} \left( 1 - \frac{\lambda_i^n}{n \mu_i} \right) \text{ exists, for all } i \in \{1, \dots, d\}. \quad (1)$$

Therefore under (1) each customer class  $i$  is in criticality with respect to pool  $i$ ,  $i \in \{1, \dots, d\}$ , i.e.,  $n \approx \mathbf{r}_i^n + \rho_i \sqrt{\mathbf{r}_i^n}$  for some constant  $\rho_i$  where  $\mathbf{r}_i^n = \lambda_i^n / \mu_i$  is the mean offered load of class  $i$  to the pool  $i$ . Note that the above criticality condition is different from those generally used in multiclass multi-pool setting (Atar 2005). This condition could be seen as a generalization to Ramanan and Reiman (2003, Assumption 4.12(3)) to the many-server setting. To elaborate we recall the generalized processor sharing (GPS) network from Ramanan and Reiman (2003). In a multiclass GPS network with  $d$  customer classes and single server the server would use a fraction  $\alpha_i$ ,  $(\alpha_1, \dots, \alpha_d) \in (0, \infty)^d$  is a given probability vector, of the total processing capacity to serve class- $i$  jobs when all the job classes are available in the system, otherwise (that is when positive number of classes are empty) any excess processing capacity, normally reserved for the job classes that are empty, is redistributed among the nonempty classes in proportion to the weight vector  $(\alpha_1, \dots, \alpha_d)$ . In this case the conventional heavy traffic condition implies that  $\lim_{n \rightarrow \infty} \sqrt{n}(\mathbf{r}_i^n/n - \alpha_i)$  exists for all  $i$ , see Ramanan and Reiman (2003, Assumption 4.12(3)). Therefore (1) can be seen as a many-server analogue to the conventional heavy-traffic condition of GPS network.

Control is given by a matrix value process  $Z$  where  $Z_{ij}$  denotes the number of class- $i$  customers getting served at station  $j$ . We note that for our model a control  $Z$  need not be work-conserving. The running cost is given by  $r(\hat{Q}^n)$  where  $r$  is a non-negative, convex function of polynomial growth and  $\hat{Q}^n$  denotes the diffusion-scaled queue length vector i.e.,  $\hat{Q}^n = (1/\sqrt{n})Q$  where  $Q$  is the queue length vector. We are interested in the cost function

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[ \int_0^T r(\hat{Q}^n(s)) ds \right].$$

The value function  $\hat{V}^n$  is defined to be the infimum of the above cost where the infimum is taken over all admissible controls. One of the main results of this paper analyze asymptotic behavior of  $\hat{V}^n$  as  $n \rightarrow \infty$ . In Theorem 2.1 we show that  $\hat{V}^n$  tends to the optimal value of an ergodic control problem governed by certain class of controlled diffusions. We also study the limiting ergodic control problem and establish the existence-uniqueness results of the value function in Theorem 3.1. It is worth mentioning that results like Theorems 2.1 and 3.1 continue to hold if one considers other types of convex running cost functions which might depend on  $\hat{Z}^n$  (see Remark 5.1). Let us denote by  $i \rightarrow j$  ( $i \nrightarrow j$ ) when class- $i$  customers can (resp., can not) access service from station  $j$ . In this article we have concentrated on the situation where  $i \rightarrow j$ , for all  $i, j$ , but it is not a necessary condition for our method to work. As noted in Remark 5.2, if we impose  $i \nrightarrow j$  for some  $i \neq j$  in the above queueing model, our results continue to hold without any major change in the arguments.

*Literature review.* Scheduling control problems have a rich and profound history in queueing literature. The main goal of these problems is to schedule the customers/jobs in a network in an optimal way. But it is not always possible to find a simple policy that is optimal. It is well known that for various queueing networks with finitely many servers  $c\mu$  policy is optimal (Atar and Biswas 2014, Biswas 2014, Mandelbaum and Stolyar 2004, van Mieghem 1995).  $c\mu$  scheduling rule prioritize the service of the job classes according to the index  $c_i\mu_i$  (higher index gets priority for receiving service) where  $c_i$  denotes the holding cost for class- $i$  and  $1/\mu_i$  denotes the mean service time of class- $i$  jobs. In case of many servers, it is shown in Atar et al. (2010, 2011) that a similar priority policy, known as  $c\mu/\theta$ -policy, that prioritize the jobs according to the index  $c_i\mu_i/\theta_i$ ,  $\theta_i$  being the abandonment rate of class- $i$ , is optimal when the queueing system asymptotic is considered under fluid scaling and the cost is given by the long run average of a linear running cost. But existence of such simple optimal priority policies are rare in Halfin-Whitt setting. In general, by Halfin-Whitt setting we mean the number of servers  $n$  and the total offered load  $\mathbf{r}$  scale like  $n \approx \mathbf{r} + \rho\sqrt{\mathbf{r}}$  for some constant  $\rho$ . See (2) below for exact formulation for our model. There are several papers devoted to the study of control problems in Halfin-Whitt regime. Atar et al. (2004), Harrison and Zeevi (2004) studied a control problem with discounted pay-off for multiclass single pool queueing network and asymptotics of the value functions are obtained in Atar et al. (2004). Later these works are generalized to multi-pool case in Atar (2005). Atar et al. (2009) considered a simplified multiclass multi-pool control problem with a discounted cost criterion where the service rates either depend on the class or the pool but not the both, and established asymptotic optimality for the scheduling policies. Under some assumptions on the holding cost, a static priority policy is shown to be optimal in Dai and Tezcan (2008, 2011) in a multiclass multi-pool queueing network where the cost function is of finite horizon type. Gurvich and Whitt (2009) studied queue-and-idleness-ratio controls, and their associated properties and staffing implications for multiclass multi-pool networks. In Arapostathis et al. (2015) the authors considered an ergodic control problem for multiclass many-server queueing network and established convergence of the value functions. Some other works that have considered ergodic control problems for queueing networks are as follows: Budhiraja et al. (2011) considers an ergodic control problem in the conventional heavy-traffic regime and establishes asymptotic optimality, Kim and Ward (2013) studies admission control problem with an ergodic cost criterion for a single class  $M/M/N+M$  queueing network. For an inverted “V” model it is shown in Armony (2005) that the fastest-server-first policy is asymptotically optimal for minimizing the steady-state expected queue length and waiting time. Ward and Armony (2013) considered a blind fair routing policy for multiclass multi-pool networks and used simulations to validate the performance of the blind fair routing policies comparing them with non-blind policies derived from the limiting diffusion control problem. Recently, ergodic control of multiclass multi-pool queueing networks is considered in Arapostathis and Pang (2016) where the authors have addressed existence and uniqueness of solutions of the HJB (Hamilton-Jacobi-Bellman) equation associated to the limiting diffusion control problem. Asymptotic optimality results for the N-network queueing model are obtained in Arapostathis and Pang (2017). Most of these above works (Arapostathis et al. 2015; Atar 2005; Atar et al. 2004, 2009; Harrison and Zeevi 2004) on many server networks consider work-conserving policies as their *admissible* controls. It is necessary to point out some key differences of the present queueing network with the one considered in Arapostathis and Pang (2016), Atar (2005). First of all the Halfin-Whitt condition above (see (1)) is different

from Atar (2005, p. 2614) and therefore, the diffusion scalings of the customer count processes are also different. Moreover, the collection of admissible controls in Atar (2005) includes a wider class of scheduling policies which are jointly work-conserving and need not follow any class-to-pool priority, whereas for the queueing model under consideration every admissible control must obey the class-to-pool priority constrain. The particular nature of our network allow us to consider an optimal ergodic control problem with a general running cost function and to obtain asymptotic optimality (Theorem 2.1) under standard assumptions on the service rates.

*Motivation and comparison.* The above model is realistic in many scenario. For instance, in a call center different service stations are designed to serve certain type of customers and they may choose to help other type of customers when one or many servers are idle in the station. It is known that cross-training of customer sales representatives in inbound call center reduces the average number of customers in queue. We refer to Iravani et al. (2007) and the references therein for a discussion on labor cross-training and its effect on the performance on queueing networks. Since it is expensive to train every sales representative in all skills it becomes important to understand the optimal cross-training structure of the agents which reduces the average number of customers in queue. Iravani et al. (2007) uses *average shortest path length* as a metric to predict the more effective cross-training structures in terms of customer waiting times. Our model is a variant of these networks. As mentioned before, we may have  $i \rightarrow j$  for some  $i \neq j$  in our queueing network which should be interpreted as the inability of station  $j$  to serve class  $i$  jobs. It is also reasonable to have class-to-pool priority when the agents in pool  $i$  are primarily trained to serve jobs of class  $i$ ,  $i \in \{1, \dots, d\}$ , and might not be efficient in serving class  $j$ ,  $i \neq j$ . It should be noted that for our queueing model we have fixed a cross-training structure and we are trying to investigate the optimal dynamic scheduling policy that will optimize the pay-off function.

Our model also bears resemblance with Generalized Processor Sharing (GPS) models from Ramanan and Reiman (2003, 2008). In fact, our model can be scaled to a single pool case where each class of customers have priority in accessing a fixed fraction of the total number of servers and they may get access to other servers, fixed for other customer classes, if the queues of other customer classes are empty. It should be observed that the multi-pool version is more general than the single pool version. For instance, it is not natural to have  $\mu_{ij} \neq \mu_i$ , for  $i \neq j$ , in the setting of a single pool with homogeneous servers, but this is not the case for a multi-pool model. Therefore we stick to the multi-pool network model. A legitimate question for these processor sharing type model is that whether the GPS type policy is optimal or not for the pay-off function considered above. Motivated by this question a similar control problem is considered in Biswas et al. (2017) for a queueing model with finitely many servers and it is shown that the value function associated to the limiting controlled diffusion model solves a non-linear Neumann boundary value problem. The solution is obtained in the viscosity sense and therefore, it is hard to extract any information about the optimal control, even for the diffusion control problem. The present work is also motivated by a similar question but for the many-server queueing network. One motivation of the present work is to get some insight about the optimal control. The motivating question here is if we allow the non-priority classes of pool  $i$  to access the servers of pool  $i$  in some fixed proportion when pool  $i$  has some free servers then such scheduling policy would be optimal or not. In the present work we characterize the value functions of the queueing model by its limit and construct a sequence of admissible policies that are asymptotically optimal. Though theoretically we can find a minimizer for the limiting HJB, it is not easy to compute it explicitly or numerically. One of the future directions is to compute the minimizer and compare its performance with the GPS type scheduling.

The methodology of this problem is not immediate from any existing work. In general, the main idea is to convert such problems to a controlled diffusion problem and analyze the corresponding Hamilton-Jacobi-Bellman (HJB) equation to extract information about the minimizing policies. All the exiting works (Arapostathis et al. 2015, Atar 2005, Atar et al. 2004, Harrison and Zeevi 2004) use work-conservative properties of the controls to come up with an action space that does not depend on the state variable. But as we mentioned above that our policies need not be work-conserving. Also there is an obvious action space that one could associate to our model (see (10)). Unfortunately, this action space depends on the state variable. In general, such action spaces are not very favorable for mathematical analysis. Existence of measurable selectors and regularity of Hamiltonian do not become obvious due to the dependency of action space on the state variable. Interestingly, for our model we could show that the structure of drift and convexity of the running cost play in favour of our analysis and we can work with such state dependent action spaces. In particular, we obtain uniform stability (Lemma 4.2) and also show that the Hamiltonian is locally Hölder continuous (Lemmas 4.3 and 4.6). Since our action space depends on state variable we need to verify that the Filipov's criterion holds (Aliprantis and Border 2006, Chapter 18) and then by using Filipov's implicit function theorem we establish existence of a measurable minimizer for the Hamiltonian. This is done in Theorem 3.1. But such a minimizer need not be continuous, in general, and one often requires a continuous minimizing selector to construct  $\varepsilon$ -optimal policies for the

value functions  $\hat{V}^n$  (see Arapostathis et al. 2015, Atar et al. 2004). With this in mind, we consider a perturbed problem where we perturb the cost by a strictly convex function and show that the perturbed hamiltonian has a unique continuous selector (Lemma 4.1). In Theorem 3.1 below we show that this continuous minimizing selector is optimal for the perturbed ergodic control problem and can be used to construct  $\varepsilon$ -optimal policies (Theorem 5.2). To summarize our contribution in this paper, we have

- considered an ergodic control problem for the  $M/M/N + M$  queueing network with labor cross-training and identified the limit of the value functions,
- solved the limiting HJB and established asymptotic optimality.

*Notations.* By  $\mathbb{R}^d$  we denote the  $d$ -dimensional Euclidean space equipped with the Euclidean norm  $|\cdot|$ . We denote by  $\mathbb{R}^{d \times d}$  the set of all  $d \times d$  real matrices and we endow this space with the usual metric. For  $a, b \in \mathbb{R}$ , we denote the maximum (minimum) of  $a$  and  $b$  as  $a \vee b$  ( $a \wedge b$ , respectively). We define  $a^+ = a \vee 0$  and  $a^- = -a \wedge 0$ . For given  $a \in \mathbb{R}$ ,  $[a]$  denotes the largest integer that is smaller or equal to  $a$ . Given a topological space  $\mathcal{X}$  and  $B \subset \mathcal{X}$ , the interior, closure, complement and boundary of  $B$  in  $\mathcal{X}$  are denoted by  $B^0$ ,  $\bar{B}$ ,  $B^c$  and  $\partial B$ , respectively.  $1_B$  is used to denote the characteristic function of the set  $B$ . By  $\mathcal{B}(\mathcal{X})$  we denote the Borel  $\sigma$ -field of  $\mathcal{X}$ . Let  $\mathcal{C}([0, \infty); \mathcal{X})$  be the set of all continuous functions from  $[0, \infty)$  to  $\mathcal{X}$ . Given a path  $f: \mathbb{R}_+ \rightarrow \mathbb{R}$ , we denote by  $\Delta f(t)$ , jump of  $f$  at time  $t$ , i.e.,  $\Delta f(t) = f(t) - f(t-)$ . We define  $\mathcal{C}^k(\mathbb{R}^d)$ ,  $k \in \mathbb{N}$ , as the set of all real valued  $k$  times continuously differentiable functions on  $\mathbb{R}^d$ . For  $\alpha \in (0, 1)$ ,  $\mathcal{C}_{\text{loc}}^{k, \alpha}(\mathbb{R}^d)$  denotes the set of all real valued  $k$ -times continuously differentiable function on  $\mathbb{R}^d$  with its  $k$ -th derivative being locally  $\alpha$ -Hölder continuous on  $\mathbb{R}^d$ . For any any domain  $D \subset \mathbb{R}^d$ ,  $\mathcal{W}^{k, p}(D)$ ,  $p \geq 1$ , denotes the set of all  $k$ -times weakly differentiable functions that is in  $L^p(D)$  and all its weak derivatives up to order  $k$  are also in  $L^p(D)$ . By  $\mathcal{W}_{\text{loc}}^{k, p}(D)$ ,  $p \geq 1$ , we denote the collection of function that are  $k$ -times weakly differentiable and all its derivatives up to order  $k$  are in  $L_{\text{loc}}^p(D)$ .  $\mathcal{C}_{\text{pol}}(\mathbb{R}_+^d)$  denotes the set of all real valued continuous functions  $f$  that have at most polynomial growth i.e.,

$$\limsup_{|x| \rightarrow 0} \frac{|f(x)|}{|x|^k} = 0, \quad \text{for some } k \geq 0.$$

For a measurable  $f$  and measure  $\mu$  we denote  $\langle f, \mu \rangle = \int f d\mu$ . Let  $\mathcal{O}(g)$  denote the space of function  $f \in \mathcal{C}(\mathbb{R}^d)$  such that

$$\sup_{x \in \mathbb{R}^d} \frac{|f(x)|}{1 + |g(x)|} < \infty.$$

By  $\mathfrak{o}(g)$  we denote the subspace of  $\mathcal{O}(g)$  containing function  $f$  satisfying

$$\limsup_{|x| \rightarrow \infty} \frac{|f(x)|}{1 + |g(x)|} = 0.$$

Infimum over empty set is regarded as  $+\infty$ .  $\kappa_1, \kappa_2, \dots$ , are deterministic positive constants whose value might change from line to line.

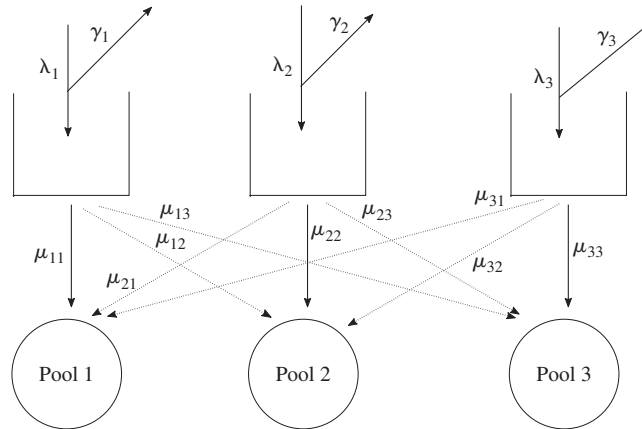
The organization of the paper is as follows. The next section introduces the setting of our model and state our main result on the convergence of the value functions. In Section 3 we formulate the limiting controlled diffusion and state our results on the ergodic control problem with state dependent action space. Section 4 obtains various results for the controlled diffusion and its HJB which are used to prove Theorem 3.1 from Section 3. Finally, in Section 5 we obtain asymptotic lower and upper bounds for the value functions.

## 2. Setting and Main Result

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a given complete probability space and all the stochastic variables introduced below are defined on it. The expectation w.r.t.  $\mathbb{P}$  is denoted by  $\mathbb{E}$ . We consider a multiclass Markovian many-server queueing system which consists of  $d$  customer classes and  $d$  server pools. Each server pool is assumed to contain  $n$  identical servers (see Figure 1).

The system buffers are assumed to have infinite capacity. Customers of class  $i \in \{1, \dots, d\}$  arrive according to a Poisson process with rate  $\lambda_i^n > 0$ . Upon arrival, customers enter the queue of their respective classes if not being processed. Customers of each class are served in the first-come-first-serve (FCFS) service discipline. Customers can abandon the system while waiting in the queue. Patience time of the customers are assumed to be exponentially distributed and class dependent. Customers of class  $i$ ,  $i \in \{1, \dots, d\}$ , renege at rate  $\gamma_i^n$ . We also assume that no customer renege while in service. Customers of class  $i$  have highest priority in accessing service from station  $i$ . A customer of class  $i$ ,  $i \in \{1, \dots, d\}$ , is allowed to access service from station  $j$ ,  $j \in \{1, \dots, d\}$ ,  $i \neq j$ ,

**Figure 1.**  $d = 3$ . Dotted lines represents access of other classes to the pool in absence of its priority class customers in the queue.



if and only if the  $j$ -th queue is empty and all the servers in the  $i$ -th pool are occupied by class- $i$  customers. By  $\mu_{ij}^n$ ,  $i, j \in \{1, \dots, d\}$ , we denote the service rate of class  $i$  at station  $j$ . We denote  $\mu_{ii}^n$  by  $\mu_i^n$  for  $i \in \{1, \dots, d\}$ . We assume that customer arrivals, service and abandonment of all classes are mutually independent.

*The Halfin-Whitt Regime.* We study this queueing model in the Halfin-Whitt regime (or the Quality-and-Efficiency-Driven (QED) regime). We consider a sequence of systems indexed by  $n$  where the arrival rates  $\lambda_i^n$  and  $n$  grows to infinity at certain rates. Let  $\mathbf{r}_i^n := \lambda_i^n / \mu_i^n$  be the mean offered load of class  $i$  customers. In the Halfin-Whitt regime, the parameters are assumed to satisfy the following: as  $n \rightarrow \infty$ ,

$$\begin{aligned} \frac{\lambda_i^n}{n} &\rightarrow \lambda_i > 0, & \mu_i^n &\rightarrow \mu_i > 0, & \mu_{ij}^n &\rightarrow \mu_{ij} \geq 0, \quad \text{for } i \neq j, \\ \gamma_i^n &\rightarrow \gamma_i > 0, & \frac{\lambda_i^n - n\lambda_i}{\sqrt{n}} &\rightarrow \hat{\lambda}_i, & \sqrt{n}(\mu_i^n - \mu_i) &\rightarrow \hat{\mu}_i, \\ \frac{\mathbf{r}_i^n}{n} &\rightarrow 1 = \frac{\lambda_i}{\mu_i}. \end{aligned} \quad (2)$$

We note that  $\mu_{ij}$ ,  $i \neq j$ , could also be 0 for some  $i \neq j$ .  $\mu_{ij} = 0$  could be understood as a situation where servers at station  $j$  are very inefficient in serving class- $i$  customers.

*State Descriptors.* Let  $X_i^n = \{X_i^n(t) : t \geq 0\}$  be the total number of class  $i$  customers in the system and  $Q_i^n = \{Q_i^n(t) : t \geq 0\}$  be the number of class  $i$  customers in the queue. By  $Z_{ij}^n$ ,  $i, j \in \{1, \dots, d\}$ , we denote the number of class  $i$  customers at the station  $j$ . As earlier we denote  $Z_{ii}^n$  by  $Z_i^n$  for  $i \in \{1, \dots, d\}$ . The following basic relationships hold for these processes: for each  $t \geq 0$ , and  $i = 1, \dots, d$ ,

$$X_i^n(t) = Q_i^n(t) + Z_i^n + \sum_{j: j \neq i} Z_{ij}^n(t), \quad Q_i^n(t) \geq 0, \quad Z_{ij}^n(t) \geq 0, \quad \text{and} \quad Z_i^n + \sum_{k: k \neq i} Z_{ki}^n(t) \leq n. \quad (3)$$

Let  $\{A_i^n, S_i^n, S_{ij}^n, R_i^n, i, j = 1, \dots, d\}$  be a collection of independent rate-1 Poisson processes. Define

$$\begin{aligned} \tilde{A}_i^n(t) &:= A_i^n(\lambda_i^n t), & \tilde{S}_i^n(t) &:= S_i^n\left(\mu_i^n \int_0^t Z_i^n(s) ds\right), \\ \tilde{S}_{ij}^n(t) &:= S_{ij}^n\left(\mu_{ij}^n \int_0^t Z_{ij}^n(s) ds\right), & \tilde{R}_i^n(t) &:= R_i^n\left(\gamma_i^n \int_0^t Q_i^n(s) ds\right). \end{aligned}$$

Then the dynamics takes the form

$$X_i^n(t) = X_i^n(0) + \tilde{A}_i^n(t) - \tilde{S}_i^n(t) - \sum_{j: j \neq i} \tilde{S}_{ij}^n(t) - \tilde{R}_i^n(t), \quad t \geq 0, \quad i = 1, \dots, d. \quad (4)$$

*Scheduling Control.* We will consider policies that are non-anticipative. We also allow preemption. Under these policies every customer class and its associated station must follow a work-conserving constrain in the following sense: for all  $i \in \{1, \dots, d\}$ ,

$$(X_i^n(t) - n)^+ = Q_i^n(t) + \sum_{j: j \neq i} Z_{ij}^n(t), \quad t \geq 0. \quad (5)$$

Combining (3) and (5) we see that

$$Z_i^n = X_i^n \wedge n, \quad 1 \leq i \leq d, \quad t \geq 0. \quad (6)$$

Therefore, when a server from station  $i$  becomes free and there are no customers of class  $i$  waiting in the queue, the server may process a customer of class  $j$ ,  $i \neq j$ . Also a customer of class  $i$  does not receive service from a server at the station  $j$ ,  $j \neq i$ , if there is an empty server at station  $i$ . Service preemption is allowed, i.e., service of a customer class can be interrupted at any time to serve some other class of customers and the original service is resumed at a later time, possibly by a server at some other station. It should be noted that a policy need not be work-conserving. For instance, it could happen that under some policy there are empty servers at station  $j$  but there could be queue of class  $i$ ,  $i \neq j$ .

Define the  $\sigma$ -fields as follows

$$\begin{aligned} \mathcal{F}_t^n &:= \sigma\{X^n(0), \tilde{A}_i^n(t), \tilde{S}_i^n(t), \tilde{S}_{ij}^n(t), \tilde{R}_i^n(t): i = 1, \dots, d, 0 \leq s \leq t\} \vee \mathcal{N}, \quad \text{and} \\ \mathcal{G}_t^n &:= \sigma\{\delta \tilde{A}_i^n(t, r), \delta \tilde{S}_i^n(t, r), \delta \tilde{S}_{ij}^n(t, r), \delta \tilde{R}_i^n(t, r): i = 1, \dots, d, r \geq 0\}, \end{aligned}$$

where

$$\begin{aligned} \delta \tilde{A}_i^n(t, r) &:= \tilde{A}_i^n(t+r) - \tilde{A}_i^n(t), \quad \delta \tilde{S}_i^n(t, r) := S_i^n\left(\mu_i^n \int_0^t Z_i^n(s) ds + \mu_i^n r\right) - \tilde{S}_i^n(t), \\ \delta \tilde{S}_{ij}^n(t, r) &:= S_{ij}^n\left(\mu_{ij}^n \int_0^t Z_{ij}^n(s) ds + \mu_{ij}^n r\right) - \tilde{S}_{ij}^n(t), \quad \delta \tilde{R}_i^n(t, r) := R_i^n\left(\gamma_i^n \int_0^t Q_i^n(s) ds + \gamma_i^n r\right) - \tilde{R}_i^n(t), \end{aligned}$$

and  $\mathcal{N}$  is the collection of all  $\mathbb{P}$ -null sets. The filtration  $\{\mathcal{F}_t^n, t \geq 0\}$  represents the information available up to time  $t$  while  $\mathcal{G}_t^n$  contains the information about future increments of the processes.

We say that a control policy is *admissible* if it satisfies (5) (or, equivalently (6)) and,

- (i)  $Z^n(t)$  is adapted to  $\mathcal{F}_t^n$ ,
- (ii)  $\mathcal{F}_t^n$  is independent of  $\mathcal{G}_t^n$  at each time  $t \geq 0$ ,
- (iii) for each  $i = 1, \dots, d$ , and  $t \geq 0$ , the process  $\delta \tilde{S}_i^n(t, \cdot)$  ( $\delta \tilde{S}_{ij}^n(t, \cdot)$ ) agrees in law with  $S_i^n(\mu_i^n \cdot)$  ( $S_{ij}^n(\mu_{ij}^n \cdot)$ ), and the process  $\delta \tilde{R}_i^n(t, \cdot)$  agrees in law with  $R_i^n(\gamma_i^n \cdot)$ .

By criterion (iii) above the increments of the processes have same distribution as the original processes in addition to being independent of  $\mathcal{F}_t^n$  (by (ii) above). We denote the set of all admissible control policies  $(Z^n, \mathcal{F}^n, \mathcal{G}^n)$  by  $\Pi^n$ .

## 2.1. Control Problem Formulation

Define the diffusion-scaled processes

$$\hat{X}^n = (\hat{X}_1^n, \dots, \hat{X}_d^n)^\top, \quad \hat{Q}^n = (\hat{Q}_1^n, \dots, \hat{Q}_d^n)^\top, \quad \text{and} \quad \hat{Z}^n = [\hat{Z}_{ij}^n],$$

by

$$\hat{X}_i^n(t) := \frac{1}{\sqrt{n}}(X_i^n(t) - n), \quad \hat{Q}_i^n(t) := \frac{1}{\sqrt{n}}Q_i^n(t), \quad \hat{Z}_{ij}^n := \frac{1}{\sqrt{n}}Z_{ij}^n, \quad i \neq j, \quad \hat{Z}_i^n(t) := \frac{1}{\sqrt{n}}(Z_i^n(t) - n) \quad (7)$$

for  $t \geq 0$ . By (4) and (6), we can express  $\hat{X}_i^n$  as

$$\begin{aligned} \hat{X}_i^n(t) &= \hat{X}_i^n(0) + \ell_i^n t + \mu_i^n \int_0^t (\hat{X}_i^n)^-(s) ds - \sum_{j: j \neq i} \mu_{ij}^n \int_0^t \hat{Z}_{ij}^n(s) ds - \gamma_i^n \int_0^t \hat{Q}_i^n(s) ds \\ &\quad + \hat{M}_{A,i}^n(t) - \hat{M}_{S,i}^n(t) - \sum_{j: j \neq i} \hat{M}_{S,ij}^n(t) - \hat{M}_{R,i}^n(t), \end{aligned} \quad (8)$$

where  $\ell^n = (\ell_1^n, \dots, \ell_d^n)^\top$  is defined as

$$\ell_i^n := \frac{1}{\sqrt{n}}(\lambda_i^n - \mu_i^n n) = \frac{\lambda_i^n - n\lambda_i}{\sqrt{n}} - \sqrt{n}(\mu_i^n - \mu_i),$$

and

$$\begin{aligned} \hat{M}_{A,i}^n(t) &:= \frac{1}{\sqrt{n}}(A_i^n(\lambda_i^n t) - \lambda_i^n t), & \hat{M}_{S,i}^n(t) &:= \frac{1}{\sqrt{n}}\left(S_i^n\left(\mu_i^n \int_0^t Z_i^n(s) ds\right) - \mu_i^n \int_0^t Z_i^n(s) ds\right), \\ \hat{M}_{S,ij}^n(t) &:= \frac{1}{\sqrt{n}}\left(S_{ij}^n\left(\mu_{ij}^n \int_0^t Z_{ij}^n(s) ds\right) - \mu_{ij}^n \int_0^t Z_{ij}^n(s) ds\right), & \hat{M}_{R,i}^n(t) &:= \frac{1}{\sqrt{n}}\left(R_i^n\left(\gamma_i^n \int_0^t Q_i^n(s) ds\right) - \gamma_i^n \int_0^t Q_i^n(s) ds\right), \end{aligned} \quad (9)$$

are square integrable martingales w.r.t. the filtration  $\{\mathcal{F}_t^n\}$ .

Note that by (2)

$$\ell_i^n \xrightarrow{n \rightarrow \infty} \ell_i := \hat{\lambda}_i - \hat{\mu}_i.$$

By  $\mathbb{R}_+^{d \times d}$  we denote the set of real matrices with non-negative entries. Define

$$\mathbf{M} := \left\{ m \in \mathbb{R}_+^{d \times d} : u_{ii} = 0, \sum_{k: k \neq i} u_{ki} \leq 1, \forall i \right\}.$$

For any  $x \in \mathbb{R}^d$ , we define

$$\mathbf{M}(x) := \left\{ u \in \mathbf{M} : \sum_{j: j \neq i} m_{ij} x_j^- \leq x_i^+, \forall i \right\}. \quad (10)$$

It is easy to see that  $\mathbf{M}(x)$  is a non-empty convex and compact subset of  $\mathbf{M}$  for all  $x \in \mathbb{R}^d$ . Also  $0 \in \mathbf{M}(x)$ , for all  $x \in \mathbb{R}^d$ . From (3), (5) and (6) we have for  $i \in \{1, \dots, d\}$ ,

$$\sum_{k: k \neq i} \hat{Z}_{ki}^n \leq (\hat{X}_i^n)^-, \quad \sum_{j: j \neq i} \hat{Z}_{ij}^n \leq (\hat{X}_i^n)^+. \quad (11)$$

Define

$$\hat{U}_{ki}^n(t) := \frac{\hat{Z}_{ki}^n(t)}{(\hat{X}_i^n(t))^-}, \quad t \geq 0,$$

where we fix  $\hat{U}_{ki}^n(t) = 0$  if  $(\hat{X}_i^n(t))^- = 0$ . We also set  $\hat{U}_{ii}^n(t) = 0$ , for all  $i \in \{1, \dots, d\}$ , and  $t \geq 0$ . Therefore using (11) we obtain,

$$\sum_{k: k \neq i} \hat{U}_{ki}^n(t) \leq 1, \quad \sum_{j: j \neq i} \hat{U}_{ij}^n(t) (\hat{X}_j^n(t))^- \leq (\hat{X}_i^n(t))^+. \quad (12)$$

Thus  $\hat{U}^n(t) \in \mathbf{M}(\hat{X}^n(t))$  for all  $t$  and  $\hat{U}^n(t)$  is  $\mathcal{F}_t^n$  adapted. Also  $\hat{U}_{ij}^n$  represents the fraction of the number of servers  $(X_j^n - n)^-$  at station  $j$  that are serving class- $i$  customers. As we show later, it is convenient to view  $\hat{U}^n(t)$  as the control.

**2.1.1. The Cost Minimization Problem.** We next introduce the running cost function for the control problem. Let  $r: \mathbb{R}_+^d \rightarrow \mathbb{R}_+$  be a given function satisfying

$$0 \leq r(x) \leq c_1(1 + |x|^m), \quad \text{for some } m \geq 1, \quad (13)$$

and some positive constant  $c_1$ . We also assume that  $r$  is convex and therefore, locally Lipschitz. For example, if we let  $h_i$ ,  $h_i \geq 0$ , be the holding cost rate for class  $i$  customers, then some of the typical running cost functions are the following:

$$r(x) = \sum_{i=1}^d h_i x_i^{m_i}, \quad \min_i m_i \geq 1.$$

These running cost functions evidently satisfy the condition in (13).

Given the initial state  $X^n(0)$  and an admissible scheduling policy  $Z^n \in \mathbb{U}^n$ , we define the diffusion-scaled cost function as

$$J(\hat{X}^n(0), \hat{Z}^n) := \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[ \int_0^T r(\hat{Q}^n(s)) ds \right], \quad (14)$$

where the running cost function  $r$  satisfies (13). Then, the associated cost minimization problem is defined by

$$\hat{V}^n(\hat{X}^n(0)) := \inf_{Z^n \in \mathbb{U}^n} J(\hat{X}^n(0), \hat{Z}^n). \quad (15)$$

We refer to  $\hat{V}^n(\hat{X}^n(0))$  as the *diffusion-scaled value function* given the initial state  $\hat{X}^n(0)$  for the  $n$ th system.

From (5) we see that for  $i \in \{1, \dots, d\}$ , and  $t \geq 0$ ,

$$\hat{Q}_i^n(t) = (\hat{X}_i^n(t))^+ - \sum_{j: j \neq i} \hat{U}_{ij}^n(t) (\hat{X}_j^n(t))^-.$$

Therefore redefining  $r$  as

$$r(x, u) = r \left( x_1^+ - \sum_{j: j \neq 1} u_{1j} x_j^-, \dots, x_d^+ - \sum_{j: j \neq d} u_{dj} x_j^- \right), \quad u \in \mathbf{M}(x), \quad (16)$$

we can rewrite the control problem as

$$\hat{V}^n(\hat{X}^n(0)) = \inf \tilde{J}(\hat{X}^n(0), \hat{U}^n),$$

where

$$\tilde{J}(\hat{X}^n(0), \hat{U}^n) := \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[ \int_0^T r(\hat{X}^n(s), \hat{U}^n(s)) ds \right], \quad (17)$$

and the infimum is taken over all admissible pairs  $(\hat{X}^n, \hat{U}^n)$  satisfying (12). Hence  $\hat{U}^n(t) \in \mathbf{M}(\hat{X}^n(t))$ , almost surely, for all  $t \geq 0$ .

For simplicity we assume that the initial condition  $\hat{X}^n(0)$  is deterministic and  $\hat{X}^n(0) \rightarrow x$ , as  $n \rightarrow \infty$ , for some  $x \in \mathbb{R}^d$ .

**2.1.2. The Limiting Controlled Diffusion Process.** As in Arapostathis et al. (2015), Atar et al. (2004), Harrison and Zeevi (2004), the analysis will be done by studying the limiting controlled diffusions. One formally deduces that, provided  $\hat{X}^n(0) \rightarrow x$ , there exists a limit  $X$  for  $\hat{X}^n$  on every finite time interval, and the limit process  $X$  is a  $d$ -dimensional diffusion process, that is,

$$dX_t = b(X_t, U_t) dt + \Sigma dW_t, \quad (18)$$

with initial condition  $X_0 = x$ . In (18) the drift  $b(x, u): \mathbb{R}^d \times \mathbf{M} \rightarrow \mathbb{R}^d$  takes the form

$$b_i(x, u) = \ell_i + \mu_i x_i^- - \sum_{j: j \neq i} \mu_{ij} u_{ij} x_j^- - \gamma_i \left( x_i^+ - \sum_{j: j \neq i} u_{ij} x_j^- \right), \quad \text{with } \ell := (\ell_1, \dots, \ell_d)^\top. \quad (19)$$

The control  $U_t$  lives in  $\mathbf{M}$  and is non-anticipative,  $W$  is a  $d$ -dimensional standard Wiener process independent of the initial condition  $X_0 = x$ , and the covariance matrix is given by

$$\Sigma \Sigma^\top = \text{diag}(2\lambda_1, \dots, 2\lambda_d).$$

A formal derivation of the drift in (19) can be obtained from (5) and (8). We also need the control to satisfy  $U_t \in \mathbf{M}(X(t))$  for all  $t \geq 0$ . We define  $q: \mathbb{R}^d \times \mathbf{M} \rightarrow \mathbb{R}^d$  as follows,

$$q_i(x, u) = x_i^+ - \sum_{j: j \neq i} u_{ij} x_j^-, \quad i \in \{1, \dots, d\}. \quad (20)$$

Thus from (19) we get that

$$b_i(x, u) = \ell_i + \mu_i x_i^- - \sum_{j: j \neq i} \mu_{ij} u_{ij} x_j^- - \gamma_i q_i(x, u). \quad (21)$$

A detailed description of Equation (18) and related results are given in Section 3.

**2.1.3. The Ergodic Control Problem for Controlled Diffusion.** Define  $\tilde{r}: \mathbb{R}_+^d \times \mathbf{M} \rightarrow \mathbb{R}_+$ , by

$$\tilde{r}(x, u) := r(q_1^+(x, u), \dots, q_d^+(x, u)).$$

We note that for  $u \in \mathbf{M}(x)$  the cost  $\tilde{r}(x, u)$  agrees with  $r(x, u)$  given by (16). In analogy with (17) we define the ergodic cost associated with the controlled diffusion process  $X$  and the running cost function  $\tilde{r}(x, u)$  as

$$J(x, U) := \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_x^U \left[ \int_0^T \tilde{r}(X_t, U_t) dt \right], \quad U \in \mathbb{U}.$$

Here  $\mathbb{U}$  denotes set of all admissible controls which are defined in Section 3. We consider the ergodic control problem

$$\varrho_*(x) = \inf_{U \in \mathbb{U}} J(x, U). \quad (22)$$

We call  $\varrho_*(x)$  the optimal value at the initial state  $x$  for the controlled diffusion process  $X$ . It is shown later that  $\varrho_*(x)$  is independent of  $x$ . A detailed treatment and related results corresponding to the ergodic control problem are given in Section 3.

Next we state the main result of this section, the proof of which can be found in Section 5.

**Theorem 2.1.** Let  $\hat{X}^n(0) \rightarrow x \in \mathbb{R}^d$ , as  $n \rightarrow \infty$ . Also assume that (2) and (13) hold where the cost function  $r$  is convex. Then

$$\lim_{n \rightarrow \infty} \hat{V}^n(\hat{X}^n(0)) = \varrho_*,$$

where  $\varrho_*(x)$  is given by (22).

Theorem 2.1 is similar in Arapostathis et al. (2015, Theorems 2.1 and 2.2). The central idea of the proof of Theorem 2.1 is same as that of Arapostathis et al. (2015). One of the main advantage of the present setting is the

stability of the system. We could directly show that for all  $n$  large the mean-empirical measures corresponding to  $n$ -th system has all polynomial moments finite under every admissible policy (see Lemma 5.1). This is not the case in Arapostathis et al. (2015) where a spatial truncation method is used to treat such difficulty. We must also note that the action space in our setting depends on the location  $x$  whereas in Arapostathis et al. (2015) the action space is a fixed compact set. Therefore we need to adopt suitable modification for this problem. As shown below the convexity property of the cost and the structure of drift  $b$  play a key role in our analysis.

### 3. An Ergodic Control Problem for Controlled Diffusions

#### 3.1. The Controlled Diffusion Model

The dynamics are modelled by controlled diffusion processes  $X = \{X(t), t \geq 0\}$  taking values in  $\mathbb{R}^d$ , and governed by the Itô stochastic differential equation

$$dX(t) = b(X(t), U(t)) dt + \Sigma dW(t), \quad (23)$$

where  $b$  is given by (21) and

$$\Sigma \Sigma^T = \text{diag}(2\lambda_1, \dots, 2\lambda_d).$$

All random processes in (23) live in a complete probability space  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ . The process  $W$  is a  $d$ -dimensional standard Brownian motion independent of the initial condition  $X_0$ .

**Definition 3.1.** A process  $U$ , taking values in  $\mathbf{M}$  and  $U(t)(\omega)$  is jointly measurable in  $(t, \omega) \in [0, \infty) \times \Omega$ , is said to be an admissible control if, there exists a strong solution  $X = \{X(t): t \geq 0\}$  satisfying (23), and,

- $U$  is non-anticipative: for  $s < t$ ,  $W_t - W_s$  is independent of

$$\mathcal{F}_s := \text{the completion of } \sigma\{X_0, U(p), W(p), p \leq s\} \text{ relative to } (\mathcal{F}, \mathbb{P}).$$

- $U(t) \in \mathbf{M}(X(t))$ , almost surely, for  $t \geq 0$ .

We let  $\mathcal{U}$  denote the set of all admissible controls. Note that the drift  $b$  is Lipschitz continuous and the diffusion matrix  $\Sigma$  non-degenerate. Since  $0 \in \mathbf{M}(x)$  for all  $x \in \mathbb{R}^d$  we see that  $U \equiv 0$  is in  $\mathcal{U}$ . Thus  $\mathcal{U}$  is non-empty. Let  $a := \Sigma \Sigma^T$ . We define the family of operators  $L^u: \mathcal{C}^2(\mathbb{R}^d) \rightarrow \mathcal{C}(\mathbb{R}^d)$ , with parameter  $u \in \mathbf{M}$ , by

$$L^u f(x) := \frac{1}{2} \sum_{i,j=1}^d a_{ij} \partial_{ij} f(x) + \sum_{i=1}^d b_i(x, u) \partial_i f(x), \quad u \in \mathbf{M}. \quad (24)$$

We refer to  $L^u$  as the *controlled extended generator* of the diffusion (23). In (24) and elsewhere in this paper we have adopted the notation  $\partial_i := \partial/\partial x_i$  and  $\partial_{ij} := \partial^2/(\partial x_i \partial x_j)$ .

A control  $U \in \mathcal{U}$  is said to be stationary Markov if for some measurable  $v: \mathbb{R}^d \rightarrow \mathbf{M}$  we have  $U(t) = v(X(t))$ . Therefore for a stationary Markov control we have  $v(X(t)) \in \mathbf{M}(X(t))$  for all  $t \geq 0$ . By  $\mathcal{U}_{\text{SM}}$  we denote the set of all stationary Markov controls.

Now we introduce relaxed controls which will be useful for our analysis. Association of relaxed controls to such control problems is useful since it extends the action space to a compact, convex set (Arapostathis et al. 2012). In our setup we show that we can not do better even we extend the controls to include relaxed controls (see Theorems 3.1 and 4.1). Moreover, relaxed control would be useful to prove asymptotic lower bounds (Theorem 5.1). By  $\mathcal{P}(\mathbf{M})$  we denote the set of all probability measures on  $\mathbf{M}$ . We can extend the drift  $b$  and the running cost  $r$  on  $\mathcal{P}(\mathbf{M})$  as follows: for  $v(du) \in \mathcal{P}(\mathbf{M})$ ,

$$b(x, v) := \int_{\mathbf{M}} b(x, u) v(du), \quad \text{and} \quad r(x, v) := \int_{\mathbf{M}} r(x, u) v(du).$$

Controls taking values in  $\mathcal{P}(\mathbf{M})$  are known as *relaxed* controls. Similarly, we can extend the definition of stationary Markov controls to measure valued processes. A stationary Markov control  $v: \mathbb{R}^d \rightarrow \mathcal{P}(\mathbf{M})$  is said to be admissible if there is a unique strong solution to (23) and  $\langle 1_{\mathbf{M}(X(t))}, v(X(t)) \rangle = 1$ , almost surely, for all  $t \geq 0$ . We continue to denote this extended class by  $\mathcal{U}_{\text{SM}}$ . We endow the space  $\mathcal{U}_{\text{SM}}$  with the following topology (Arapostathis et al. 2012, Chapter 2.2.4):  $v_n \rightarrow v$  in  $\mathcal{U}_{\text{SM}}$  if and only if

$$\int_{\mathbb{R}^d} f(x) \int_{\mathbf{M}} g(x, u) v_n(du | x) dx \xrightarrow{n \rightarrow \infty} \int_{\mathbb{R}^d} f(x) \int_{\mathbf{M}} g(x, u) v(du | x) dx$$

for all  $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$  and  $g \in \mathcal{C}_b(\mathbb{R}^d \times \mathbf{M})$ .

**Proposition 3.1.** The space  $\mathcal{U}_{\text{SM}}$  under the above mentioned topology is compact whenever the initial condition  $X(0)$  is a fixed point.

**Proof.** Let  $X(0) = x$ . Let  $\{v_n\} \in \mathbb{U}_{\text{SM}}$  be a sequence of stationary Markov controls. Then  $(X_n(t), v_n(X_n(t)))$  satisfies (23) and

$$\langle 1_{M(X_n(t))}, v_n(X_n(t)) \rangle = 1, \quad \text{a.s., for } t \geq 0. \quad (25)$$

Now from Arapostathis et al. (2012, Section 2.4) there exists a measurable  $v: \mathbb{R}^d \rightarrow \mathcal{P}(\mathbf{M})$  such that  $v_n \rightarrow v$  as  $n \rightarrow \infty$  along some subsequence in the topology of Markov control. We continue to denote the subsequence by  $v_n$ . Now consider the strong solution  $X(t)$  corresponding to the Markov control  $v$ . Existence of the unique strong solution is assured by Arapostathis et al. (2012, Theorem 2.2.12). Moreover,  $X_n \Rightarrow X$  as  $n \rightarrow \infty$ , in  $\mathcal{C}([0, \infty), \mathbb{R}^d)$ . A similar the argument as in Arapostathis et al. (2012, Lemma 2.4.2) shows that for any  $t > 0$ ,

$$\|p_n(t, x, \cdot) - p(t, x, \cdot)\|_{L^1(\mathbb{R}^d)} \rightarrow 0, \quad \text{as } n \rightarrow \infty, \quad (26)$$

where  $p_n, p$  denote the transition density of  $X_n, X$  respectively, at time  $t$  starting from  $x$ . Also the transition densities are locally bounded. Now observe that  $(x, u) \mapsto 1_{M^c(x)}(u)$  is a lower-semicontinuous function. This fact follows from the definition of  $\mathbf{M}(x)$ . Hence there exists a sequence of bounded, continuous function  $g_k: \mathbb{R}^d \times \mathbf{M} \rightarrow \mathbb{R}$  such that (see for example, Krantz 2001, Proposition 2.1.2)

$$g_k(x, u) \nearrow 1_{M^c(x)}(u), \quad \text{pointwise, as } k \rightarrow \infty.$$

Let  $\phi_k$  be any smooth, non-negative function taking values in  $[0, 1]$  with support in  $\mathbb{B}_k(0)$ . We choose  $\phi_k$  to satisfy  $\phi_k \nearrow 1$ , as  $k \rightarrow \infty$ . Then from the convergence criterion of Markov controls we get for all  $k \geq 1$ ,

$$\begin{aligned} & \int_{\mathbb{R}^d} p(t, x, y) \phi_k(y) \int_M g_k(y, u) v(du | y) dy \\ & \leq \int_{\mathbb{R}^d} p(t, x, y) \phi_k(y) \int_M g_k(y, u) (v(du | y) - v_n(du | y)) \\ & \quad + \sup_{(y, u) \in \mathbb{R}^d \times M} |g_k(y, u)| \|p_n(t, x, \cdot) - p(t, x, \cdot)\|_{L^1(\mathbb{R}^d)} + \int_{\mathbb{R}^d} p_n(t, x, y) \phi_k(y) \int_M g_k(y, u) v_n(du | y) \\ & \leq \int_{\mathbb{R}^d} p(t, x, y) \phi_k(y) \int_M g_k(y, u) (v(du | y) - v_n(du | y)) \\ & \quad + \sup_{(y, u) \in \mathbb{R}^d \times M} |g_k(y, u)| \|p_n(t, x, \cdot) - p(t, x, \cdot)\|_{L^1(\mathbb{R}^d)} + \int_{\mathbb{R}^d} p_n(t, x, y) \int_M 1_{M^c(y)}(u) v_n(du | y) dy \rightarrow 0, \quad \text{as } n \rightarrow \infty, \end{aligned}$$

where in the last line we use (25), (26). Therefore letting  $k \rightarrow \infty$  we have for  $t > 0$  that

$$\mathbb{E} \left[ \int_M 1_{M^c(X(t))} v(du | X(t)) \right] = 0.$$

This proves that for every  $t > 0$ , support of  $v(du | X(t))$  lies in  $M(X(t))$  almost surely. Now define  $v(x) = 0$  to make sure  $v(X(0)) \in M(X(0))$ . This completes the proof.  $\square$

It is well known Arapostathis et al. (2012, Theorem 2.2.12) that for every Markov control in  $\mathbb{U}_{\text{SM}}$  there is a unique strong solution to (23) which is also a strong Markov process. A stationary control is called stable if the associated Markov process  $X$  is positive recurrent.

We recall the cost function  $\tilde{r}$  from previous section where

$$\tilde{r}(x, u) = r(q_1^+(x, u), \dots, q_d^+(x, u))$$

and  $r$  is a convex function that satisfies (13). Define

$$\|u\| := \sum_{i,j=1}^d |u_{ij}|^2, \quad \text{where } u \in \mathbf{M}.$$

For  $\varepsilon \in (0, 1)$ , we consider the following perturbed cost function

$$\tilde{r}_\varepsilon(x, u) := \tilde{r}(x, u) + \varepsilon \|u\|.$$

Since  $\|\cdot\|$  is strictly convex in  $u$ , we have  $\tilde{r}_\varepsilon(x, \cdot)$  strictly convex on  $\mathbf{M}(x)$  for every  $x \in \mathbb{R}^d$ . For  $U \in \mathbb{U}$ , we define

$$J(x, U) = \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[ \int_0^T \tilde{r}(X(s), U(s)) ds \right], \quad J_\varepsilon(x, U) = \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[ \int_0^T \tilde{r}_\varepsilon(X(s), U(s)) ds \right].$$

Therefore we have two value functions given by

$$\varrho_* := \inf_{U \in \mathbb{U}} J(x, U), \quad \varrho_\varepsilon := \inf_{U \in \mathbb{U}} J_\varepsilon(x, U). \quad (27)$$

We have suppressed the dependency of  $x$  from the value function as it is shown below that the value functions do not depend on  $x$ . Our main result of this section is the following.

**Theorem 3.1.** *There exist  $V_\varepsilon, V \in \mathcal{C}^2(\mathbb{R}^d)$  satisfying the following Hamilton-Jacobi-Bellman (HJB) equations,*

$$\min_{u \in M(x)} (L^u V_\varepsilon(x) + \tilde{r}_\varepsilon(x, u)) = \varrho_\varepsilon, \quad \text{for } \varepsilon \in (0, 1), \quad (28)$$

$$\min_{u \in M(x)} (L^u V(x) + \tilde{r}(x, u)) = \varrho_*, \quad (29)$$

such that

- (i) for each  $\varepsilon \in (0, 1)$ , there is a unique continuous selector of (28) which is also optimal for  $\varrho_\varepsilon$ ;
- (ii) if a pair  $(\bar{V}_\varepsilon, \bar{\varrho}_\varepsilon) \in \mathcal{C}^2(\mathbb{R}^d) \cap \mathcal{C}_{\text{pol}}(\mathbb{R}^d) \times \mathbb{R}$  satisfies (28) then we have  $(\bar{V}_\varepsilon, \bar{\varrho}_\varepsilon) = (V_\varepsilon, \varrho_\varepsilon)$ ;
- (iii) as  $\varepsilon \rightarrow 0$ ,  $V_\varepsilon \rightarrow V$  in  $\mathcal{W}_{\text{loc}}^{2,p}(\mathbb{R}^d)$ ,  $p \geq 1$ ;
- (iv) every measurable selector of (29) is optimal for  $\varrho_*$ ;
- (v) if a pair  $(\bar{V}, \bar{\varrho}) \in \mathcal{C}^2(\mathbb{R}^d) \cap \mathcal{C}_{\text{pol}}(\mathbb{R}^d) \times \mathbb{R}$  satisfies (29) then we have  $(\bar{V}, \bar{\varrho}) = (V, \varrho_*)$ .

It is not hard to see that  $\varrho_\varepsilon \searrow \varrho_*$ , as  $\varepsilon \rightarrow 0$  (the difference between the running costs is the order of  $\varepsilon$ ). Therefore using Theorem 3.1 we can find  $\varepsilon$ -optimal controls (in fact, continuous Markov controls) for  $\varrho_*$ . Continuity property of the Markov controls plays a key role in the construction of  $\varepsilon$ -optimal admissible controls for the queueing models and in obtaining asymptotic upper bound for Theorem 2.1. Results similar to Theorem 3.1 are also obtained in Arapostathis et al. (2015) for a fixed action space that does not depend on  $x$ . Therefore the results of Arapostathis et al. (2015) do not directly apply here. Because of the state dependency of the action space we need to put extra effort to establish regularity properties of the value function. Also, finding a measurable minimizing selector of (29) becomes less obvious. On the other hand, we have uniform stability (see Lemma 4.2), an advantage compared to Arapostathis et al. (2015), which help us in proving Theorem 3.1. It is shown that the  $\varepsilon$ -perturbed Hamiltonian in (28) has certain regularity properties (Lemma 4.3) which together with the strict convexity of the perturbed cost  $\tilde{r}_\varepsilon$  help us in finding a unique continuous minimizing selector (Lemma 4.1). Using these properties we characterize the discounted value function  $V_\varepsilon^\alpha$  with the running cost  $\tilde{r}_\varepsilon$  in Theorem 4.2. Then using uniform stability and the Sobolev estimates we show that the scaled value function  $\bar{V}_\varepsilon^\alpha := V_\varepsilon^\alpha - V_\varepsilon^\alpha(0)$  converges to a limit  $V_\varepsilon$ , as  $\alpha \rightarrow 0$ , that solves (28). A similar argument is also used to justify the passage of limit in  $V_\varepsilon \rightarrow V$  as  $\varepsilon \rightarrow 0$ . The detailed proof of Theorem 3.1 is given in Section 4.

## 4. Uniform Stability and Related Results

The goal of this section is to prove Theorem 3.1 and obtain related estimates. In what follows, we use several standard results from the theory of elliptic PDE's without explicitly mentioning its reference. Those results can be found in Gilbarg and Trudinger (1983), Arapostathis et al. (2012, appendix). For instance, the following is used in several places: if a function  $\psi \in \mathcal{W}_{\text{loc}}^{2,p}(\mathbb{B})$ ,  $p > d$ ,  $\mathbb{B}$  is an open set, satisfies  $\frac{1}{2} \sum_{ij} a_{ij} \partial_{ij} \psi(x) = f(x)$  for some function  $f \in \mathcal{C}^\alpha(\mathbb{B})$ ,  $\alpha \in (0, 1)$ , then  $\psi \in \mathcal{C}_{\text{loc}}^{2,\beta}(\mathbb{B})$  for some  $\beta > 0$ .

We start by proving continuity property of the selector. Define a map  $\varphi: \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbf{M}$  as follows.

$$\varphi(x, p) := \arg \min_{u \in M(x)} \{b(x, u) \cdot p + \tilde{r}_\varepsilon(x, u)\}.$$

Since  $\mathbf{M}(x)$  is convex and compact for every  $x$  and  $\tilde{r}_\varepsilon(x, \cdot)$  is strictly convex on  $\mathbf{M}(x)$  we see that  $\varphi$  is well defined.

**Lemma 4.1.** *For every  $\varepsilon \in (0, 1)$ , the function  $\varphi$  defined above is continuous.*

**Proof.** Consider a point  $(x, p) \in \mathbb{R}^d \times \mathbb{R}^d$ . Let  $\varphi(x, p) = u \in \mathbf{M}(x)$ . We claim that if  $x_i^- = 0$  for some  $i \in \{1, \dots, d\}$ , then

$$u_{ki} = 0, \quad \text{for all } k \neq i. \quad (30)$$

Suppose (30) is not true and  $u_{ki} > 0$  for some  $k \neq i$ . We define  $\bar{u} \in \mathbf{M}$  as follows,

$$\bar{u}_{lm} = \begin{cases} u_{j_1 j_2} & \text{if } j_1 \neq k, \text{ or } j_2 \neq i, \\ 0 & \text{if } j_1 = k, \text{ and } j_2 = i. \end{cases}$$

It is easy to check that  $\bar{u} \in \mathbf{M}(x)$ . We also have

$$b(x, u) \cdot p + \tilde{r}_\varepsilon(x, u) > b(x, \bar{u}) \cdot p + \tilde{r}_\varepsilon(x, \bar{u}).$$

which contradicts the fact that  $u \in \arg \min_{u \in \mathbf{M}(x)} \{b(x, u) \cdot p + \tilde{r}_\varepsilon(x, u)\}$  and this proves (30). Let  $(x_n, p_n) \rightarrow (x, p)$ , as  $n \rightarrow \infty$ . We show that  $\varphi(x_n, p_n) \rightarrow \varphi(x, p)$ , as  $n \rightarrow \infty$ . Let

$$u_n := \varphi(x_n, p_n) \in \arg \min_{u \in \mathbf{M}(x_n)} \{b(x_n, u) \cdot p_n + \tilde{r}_\varepsilon(x_n, u)\}, \quad u := \varphi(x, p) \in \arg \min_{u \in \mathbf{M}(x)} \{b(x, u) \cdot p + \tilde{r}_\varepsilon(x, u)\}.$$

Since the sequence  $\{u_n\}$  is bounded we can assume that  $u_n \rightarrow m_0 \in \mathbf{M}$ . It is also easy to see that  $m_0 \in \mathbf{M}(x)$ . We need to show that  $m_0 = u$ . Given  $u$  as above we find  $m_n \in \mathbf{M}(x_n)$  such that  $m_n \rightarrow u$ , as  $n \rightarrow \infty$ . Construction of  $m_n$  is done in following two cases.

Case 1: Let  $u = 0$ . Therefore we let  $m_n = 0 \in \mathbf{M}(x_n)$  for all  $n$ .

Case 2: Let  $u \neq 0$ . Then using (30) we see that whenever  $u_{ki} > 0$  for some  $k$  and  $i \neq k$  we have  $x_i^- > 0$ . Let  $\mathbf{I}(x) = \{i: u_{ki} > 0 \text{ for some } k \neq i\}$ . Therefore  $(x_n)_i^- > 0$  for all  $i \in \mathbf{I}(x)$  and large  $n$ . Define for large  $n$ ,

$$\delta_{ki}^n := \frac{2|x_n - x|}{(x_n)_i^-}, \quad \text{whenever } u_{ki} > 0.$$

We set  $\delta_{ki}^n = 0$  otherwise. Define  $(m_n)_{ki} = u_{ki} - \delta_{ki}^n$  for all  $k \neq i$ . Since  $\delta^n \rightarrow 0$  as  $n \rightarrow \infty$ , we have  $m_n \in \mathbf{M}$  for all large  $n$ . Now we show that  $m_n \in \mathbf{M}(x_n)$  for all large  $n$ . To do this we note that for any  $i \in \{1, \dots, d\}$ ,

$$\sum_{j: j \neq i} (m_n)_{ij} (x_n)_j^- = \sum_{j: u_{ij} > 0} (u_{ij} - \delta_{ij}^n) \cdot (x_n)_j^- = \sum_{j: u_{ij} > 0} u_{ij} \cdot (x_n)_j^- - |x_n - x| \sum_{j: u_{ij} > 0} 2.$$

If the set  $\{j: u_{ij} > 0\}$  is empty then the rhs of the above display is less than  $(x_n)_i^+$ . Otherwise we get

$$\sum_{j: j \neq i} (m_n)_{ij} (x_n)_j^- \leq \sum_{j: u_{ij} > 0} u_{ij} x_j^- + |x - x_n| \sum_{j: u_{ij} > 0} (u_{ij} - 2) \leq x_i^+ - |x - x_n| \sum_{j: u_{ij} > 0} 1 \leq (x_n)_i^+ + |x - x_n| \left(1 - \sum_{j: u_{ij} > 0} 1\right) \leq (x_n)_i^+,$$

where in the second inequality we use the fact that  $u \in \mathbf{M}(x)$  and  $u_{ij} \leq 1$ . This proves that  $m_n \in \mathbf{M}(x_n)$  for all large  $n$ . Hence using the definition we have

$$b(x_n, u_n) \cdot p_n + \tilde{r}_\varepsilon(x_n, u_n) \leq b(x_n, m_n) \cdot p_n + \tilde{r}_\varepsilon(x_n, m_n),$$

and letting  $n \rightarrow \infty$ , we get

$$b(x_n, m_0) \cdot p + \tilde{r}_\varepsilon(x, m_0) \leq b(x, u) \cdot p + \tilde{r}_\varepsilon(x, u).$$

Therefore  $m_0 \in \mathbf{M}(x)$  is also a minimizer in  $\mathbf{M}(x)$ . By uniqueness property of the minimizer in  $\mathbf{M}(x)$  we get  $m_0 = u$ . Hence the proof.  $\square$

Let  $\mathcal{V} \in \mathcal{C}^2(\mathbb{R}^d)$  be such that  $\mathcal{V}(x) = |x|^k$  for  $|x| \geq 1$  where  $k \geq 1$ . In fact, we can take  $\mathcal{V}$  to be non-negative. Also define

$$h(x) := |x|^k, \quad \text{for } x \in \mathbb{R}^d.$$

Following lemma establishes uniform stability of the our system.

**Lemma 4.2.** *There exists a positive constants  $c_5, c_6$ , depending on  $k$ , such that*

$$\sup_{u \in \mathbf{M}(x)} L^u \mathcal{V}(x) \leq c_5 - c_6 h(x), \quad \text{for all } x \in \mathbb{R}^d. \quad (31)$$

**Proof.** Recall from (21) that for  $u \in \mathbf{M}(x)$ ,

$$b_i(x, u) = \ell_i + \mu_i x_i^- - \sum_{j: j \neq i} \mu_{ij} u_{ij} x_j^- - \gamma_i q_i(x, u), \quad q_i(x, u) = x_i^+ - \sum_{j: j \neq i} u_{ij} x_j^-.$$

For  $x \in \mathbb{R}^d$ , and  $u \in \mathbf{M}(x)$ , we have  $q_i(x, u) \geq 0$ , and

$$\sum_{i=1}^d |x_i| = \sum_{i=1}^d (x_i^+ + x_i^-) \leq \sum_{i=1}^d \left( q_i(x, u) + \sum_{j: j \neq i} u_{ij} x_j^- + x_i^- \right) \leq \sum_{i=1}^d (q_i(x, u) + d x_i^-),$$

where in the last line above we use the fact that  $u_{ij} \leq 1$ . Hence

$$|x|^2 \leq d \sum_{i=1}^d (q_i(x, u) + d |x_i^-|)^2 \leq 2d \sum_{i=1}^d ((q_i(x, u))^2 + d^2 |x_i^-|^2). \quad (32)$$

We note that for any  $u \in \mathbf{M}(x)$ , we have

$$q_i(x, u) \cdot x_i = q_i(x, u) \cdot x_i^+ \geq q_i^2(x, u).$$

Thus for  $u \in \mathbf{M}(x)$  and  $|x| \geq 1$ , we obtain

$$\begin{aligned} \sum_{i=1}^d b_i(x, u) \cdot \partial_i \mathcal{V}(x) &\leq k|x|^{k-2} \left( \sum_{i=1}^d |\ell_i| |x_i| - \mu_i |x_i^-|^2 - x_i \sum_{j:j \neq i} \mu_{ij} u_{ij} x_j^- - \gamma_i q_i(x, u) x_i \right) \\ &\leq k|x|^{k-2} \left( \sum_{i=1}^d |\ell_i| |x_i| - \mu_i |x_i^-|^2 - \gamma_i q_i^2(x, u) \right) \\ &\leq \kappa_1 |x|^{k-1} - \kappa_2 |x|^k, \end{aligned} \quad (33)$$

for some constants  $\kappa_1, \kappa_2 > 0$ , where in the second inequality we use the fact that

$$x_i \sum_{j:j \neq i} \mu_{ij} u_{ij} x_j^- = x_i^+ \sum_{j:j \neq i} \mu_{ij} u_{ij} x_j^- \geq 0,$$

for  $u \in \mathbf{M}(x)$ , and in the third inequality we use (32). Now the proof can be seen using (33).  $\square$

Next we discuss a convex analytic approach that assures the existence of an optimal control. To do this we introduce some more notations. For  $U \in \mathbb{U}$ , we define

$$\varrho_U(x) := \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_x^U \left[ \int_0^T \tilde{r}(X(t), U(t)) dt \right].$$

We use the notation  $\mathbb{E}^U[\cdot]$  to express the dependency in  $U$ . For  $\beta > 0$ , we define

$$\mathbb{U}^\beta := \{U \in \mathbb{U}: \varrho_U(x) \leq \beta \text{ for some } x \in \mathbb{R}^d\}.$$

Let  $\mathbb{U}_{\text{SM}}^\beta = \mathbb{U}_{\text{SM}} \cap \mathbb{U}^\beta$ . Denote

$$\varrho_* := \inf\{\beta > 0: \mathbb{U}^\beta \neq \emptyset\}, \quad \hat{\varrho}_* := \inf\{\beta > 0: \mathbb{U}_{\text{SM}}^\beta \neq \emptyset\}, \quad \tilde{\varrho}_* := \inf\{\pi(\tilde{r}): \pi \in \mathcal{G}\},$$

where

$$\pi(\tilde{r}) := \int_{\mathbb{R}^d \times \mathcal{M}} \tilde{r}(x, u) \pi(dx, du),$$

and,

$$\mathcal{G} := \left\{ \pi \in \mathcal{P}(\mathbb{R}^d \times \mathbf{M}): \int_{\mathbb{R}^d \times \mathcal{M}} L^u f(x) \pi(dx, du) = 0, \text{ for all } f \in \mathcal{C}_c^\infty(\mathbb{R}^d), \text{ and } \int_{\mathbb{R}^d \times \mathcal{M}} 1_{M(x)}(u) \pi(dx, du) = 1 \right\}. \quad (34)$$

In what follows, we denote by  $\tau(A)$  the *first exit time* of a process  $\{X_t, t \in \mathbb{R}_+\}$  from a set  $A \subset \mathbb{R}^d$ , defined by

$$\tau(A) := \inf\{t > 0: X_t \notin A\}.$$

The open ball of radius  $R$  in  $\mathbb{R}^d$ , centered at the origin, is denoted by  $\mathbb{B}_R$ , and we let  $\tau_R := \tau(\mathbb{B}_R)$ , and  $\tilde{\tau}_R := \tau(\mathbb{B}_R^c)$ .

**Theorem 4.1.** *We have*

- (a)  $\varrho_* = \hat{\varrho}_* = \tilde{\varrho}_*$ .
- (b) *There exists  $v \in \mathbb{U}_{\text{SM}}$  such that  $\varrho_v = \varrho_*$ .*

**Proof.** Let  $U \in \mathbb{U}$ . Applying Itô's formula, it follows from (31) that

$$\frac{1}{T} (\mathbb{E}_x^U [\mathcal{V}(X(\tau_R \wedge T))] - \mathcal{V}(x)) \leq c_5 - c_6 \frac{1}{T} \mathbb{E}_x^U \left[ \int_0^{\tau_R \wedge T} h(X(s)) ds \right].$$

Letting  $R \rightarrow \infty$ , and then  $T \rightarrow \infty$ , we see that

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_x^U \left[ \int_0^T h(X(s)) ds \right] < \infty.$$

Since  $h$  is inf-compact, the above display implies that the *mean-empirical measures*  $\{\zeta_{x,t}: t > 0\}$ , defined by

$$\zeta_{x,t}(A \times B) := \frac{1}{t} \mathbb{E}_x \left[ \int_0^t 1_{A \times B}(X(s), U(s)) ds \right],$$

are tight. Since

$$\int_{\mathbb{R}^d \times M} 1_{M^c(x)}(u) \zeta_{x,t}(dx, du) = 0,$$

and  $(x, u) \mapsto 1_{M^c(x)}(u)$  is lower-semicontinuous, it is easy to see that every sub sequential limit of  $\{\zeta_{x,t}: t > 0\}$ , as  $t \rightarrow \infty$ , lies in  $\mathcal{G}$ . Also if  $\pi$  is one of the limits of  $\{\zeta_{x,t}: t > 0\}$ , we get

$$\pi(\tilde{r}) \leq \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_x^U \left[ \int_0^T \tilde{r}(X(s), U(s)) ds \right].$$

This shows that  $\tilde{\varrho}_* \leq \varrho_*$ . Let  $\pi \in \mathcal{G}$ . Using disintegration of measure we write  $\pi(dx, du) = v(du | x) \mu_v(dx)$ . Therefore by definition, we have

$$\int_{\mathbb{R}^d} L^v f(x) \mu_v(dx) = \int_{\mathbb{R}^d} \int_M L^u f(x) v(du | x) \mu_v(dx) = 0, \quad \forall f \in \mathcal{C}_c^\infty(\mathbb{R}^d).$$

Hence applying Arapostathis et al. (2012, Theorem 2.6.16) we see that  $\mu_v(dx)$  has locally strictly positive density. In particular,  $\mu_v(dx)$  is mutually absolutely continuous with respect to the Lebesgue measure on  $\mathbb{R}^d$ . Combining these observations with the fact that

$$\int_{\mathbb{R}^d \times M} 1_{M^c(x)}(u) \pi(dx, du) = \int_{\mathbb{R}^d \times M} v(\mathbf{M}^c(x) | x) \mu_v(dx) = 0,$$

we conclude that  $v(\mathbf{M}^c(x) | x) = 0$  almost surely with respect to the Lebesgue measure on  $\mathbb{R}^d$ . Since  $v(\mathbf{M}^c(x) | x)$  is Borel-measurable we can modify  $v$  on a Borel set of measure 0 so that  $v(\mathbf{M}^c(x) | x) = 0$  holds everywhere. Hence the stationary solution  $X(t)$  corresponds to the Markov control  $v(\cdot | x)$  would satisfy  $v(\mathbf{M}(X(t)) | X(t)) = 1$  almost surely. Thus  $v(\cdot | x)$  is an admissible Markov control. By ergodic theorem Yosida (1980), Arapostathis et al. (2012, Theorem 1.5.18) it is known that

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_x^v \left[ \int_0^T \tilde{r}(X(s), v(X(s))) ds \right] \xrightarrow{n \rightarrow \infty} \pi(\tilde{r}), \quad \text{for almost every } x \in \mathbb{R}^d.$$

Note that if  $\tilde{r}(X(s), v(X(s)))$  is continuous we could use weak convergence of mean-empirical measures, corresponding to  $X$ , to justify the above limit. But  $v$  need not be continuous, in general. So we use ergodic theorem to pass the limit. Thus  $\hat{\varrho}_* \leq \tilde{\varrho}_*$ . But by definition  $\varrho_* \leq \hat{\varrho}_*$ . Thus we have  $\varrho_* = \hat{\varrho}_* = \tilde{\varrho}_*$ . This proves (a).

To prove (b), we consider a sequence  $\pi_n \in \mathcal{G}$  along which the infimum is achieved. Applying Lemma 4.2 we obtain that the measures  $\{\pi_n\}$  are tight and thus it has a convergent subsequence. Let  $\pi$  be one of the subsequential limits and  $\pi_n \rightarrow \pi$ , after relabelling, as  $n \rightarrow \infty$ . Lower-semicontinuity of  $(x, u) \mapsto 1_{M^c(x)}(u)$  implies that  $\pi \in \mathcal{G}$ . Moreover,

$$\pi(\tilde{r}) \leq \liminf_{n \rightarrow \infty} \pi_n(\tilde{r}) = \varrho_*.$$

Thus the infimum is achieved at  $\pi$  and the Markov control is obtained by disintegrating  $\pi(dx, du) = v(du | x) \mu_v(dx)$  where  $v(du | \cdot)$  is our required Markov control.  $\square$

**Remark 4.1.** We observe that the arguments of Theorem 4.1 continue to hold if we replace  $\tilde{r}$  by  $\tilde{r}_\varepsilon$  and  $\varrho$  by  $\varrho_\varepsilon$ . The above theorem also establishes independence of  $\varrho$  (and  $\varrho_\varepsilon$ ) from  $x$ .

The above existence result of optimal Markov control is purely analytic and it does not provide any characterization of the optimal controls. To prove Theorem 2.1 we need to find an optimal control with some regularity properties such as continuity. But the above result does not say anything about the regularity properties of optimal Markov controls. Therefore we will analyze the associated HJB to extract more information about the optimal Markov controls.

For  $\alpha \in (0, 1)$ , we define

$$V_\varepsilon^\alpha(x) := \inf_{U \in \mathbb{U}} \mathbb{E}_x \left[ \int_0^\infty e^{-\alpha s} \tilde{r}_\varepsilon(X(s), U(s)) ds \right]. \quad (35)$$

The following result characterizes the  $\alpha$ -discounted problem.

**Theorem 4.2.** For every  $\varepsilon \in (0, 1)$ ,  $V_\varepsilon^\alpha \in \mathcal{C}^2(\mathbb{R}^d)$  satisfies the HJB

$$\min_{u \in M(x)} (L^u V_\varepsilon^\alpha(x) + \tilde{r}_\varepsilon(x, u)) - \alpha V_\varepsilon^\alpha = 0. \quad (36)$$

Moreover, the unique minimizing selector is an optimal Markov control for (35).

Before we prove the theorem we need to establish some regularity properties of the Hamiltonian. To do so we define,

$$H_\varepsilon(x, p) := \min_{u \in M(x)} \{p \cdot b(x, u) + \tilde{r}_\varepsilon(x, u)\}.$$

**Lemma 4.3.** Let  $\mathcal{K} \subset \mathbb{R}^d$  be compact. Then for any  $R > 0$ , there exists a constant  $c = c(\mathcal{K}, R, \varepsilon)$ , depending on  $\mathcal{K}, R$  and  $\varepsilon$ , such that

$$|H_\varepsilon(x, p) - H_\varepsilon(y, q)| \leq c(|x - y|^{1/4} + |p - q|), \quad \text{for all } x, y \in \mathcal{K}, \text{ and } p, q \in \mathbb{B}_R(0).$$

**Proof.** We note that for any  $x \in \mathcal{K}$ , we have

$$|H_\varepsilon(x, p) - H_\varepsilon(x, q)| \leq \sup_{(x, u) \in \mathcal{K} \times M} |b(x, u)| |p - q|.$$

Therefore it is enough to show that for any  $x, y \in \mathcal{K}$ ,  $q \in \mathbb{B}_R(0)$ ,

$$H_\varepsilon(x, q) - H_\varepsilon(y, q) \leq c(\mathcal{K}, R, \varepsilon) |x - y|^{1/4}. \quad (37)$$

Since  $b$  and  $r$  are locally Lipschitz, we have a constant  $\kappa > 0$ , depending on  $\mathcal{K}$ , such that

$$|b(x, u) - b(x, \bar{u})| + |\tilde{r}(x, u) - \tilde{r}(x, \bar{u})| \leq \kappa \sum_{i \neq j} |u_{ij} x_j^- - \bar{u}_{ij} x_j^-|, \quad \text{for } u, \bar{u} \in \mathbf{M}(x), \quad x \in \mathcal{K}. \quad (38)$$

Let  $u \in \mathbf{M}(y)$  be a minimizer of

$$\{q \cdot b(y, u') + \tilde{r}_\varepsilon(y, u')\},$$

in  $\mathbf{M}(y)$ . In fact, this is the unique minimizer because of the strict convexity of the functional. Define  $\theta := [(2\kappa/\varepsilon)(|q| + 1)]^{1/2} \vee d$ . We claim that

$$\text{for any } \delta \in (0, 1/\theta), \text{ there is no } u_{ij} > \theta\delta, \text{ if } y_j^- < \delta^2. \quad (39)$$

The above implies that either  $u_{ij} \leq \theta\delta$  or  $\sum_{j: u_{ij} > \theta\delta} y_j^- \geq \delta^2 \sum_{j: u_{ij} > \theta\delta} 1$ . Supposing the contrary, we assume that (39) is not true i.e., there exists  $i, j, i \neq j$ , such that

$$u_{ij} \in (\delta\theta, 1], \quad \text{and} \quad y_j^- < \delta^2.$$

We define  $\bar{u} \in \mathbf{M}(y)$  as follows

$$\bar{u}_{j_1 j_2} = \begin{cases} u_{j_1 j_2} & \text{if } j_1 \neq i, \text{ or } j_2 \neq j, \\ 0 & \text{otherwise.} \end{cases}$$

Then using (38) we obtain

$$\{q \cdot b(y, u) + \tilde{r}_\varepsilon(y, u)\} - \{q \cdot b(y, \bar{u}) + \tilde{r}_\varepsilon(y, \bar{u})\} \geq -\kappa |q| \delta^2 - \kappa \delta^2 + \varepsilon |u_{ij}|^2 \geq -(|q| + 1) \kappa \delta^2 + \varepsilon \theta^2 \delta^2 > 0.$$

This contradicts that fact that  $u \in \mathbf{M}(y)$  is a minimizer. This proves the Claim (39).

Now we proceed to prove (37). We note that it is enough to show (37) for  $|x - y| < 1/\theta$ . Let  $|x - y| < 1/\theta$ . We define  $\mathcal{A}(y) \subset \{1, \dots, d\}$  as follows.

$$\mathcal{A}(y) := \left\{ i \in \{1, \dots, d\} : \sum_{j: u_{ij} > \theta |x - y|^{1/4}} y_j^- \geq |x - y|^{1/2} \right\}.$$

Note that  $\mathcal{A}(y)$  could be empty. Observe that for any  $i \in \mathcal{A}^c$  we either have  $u_{ij} \leq \theta|x - y|^{1/4}$  for all  $j \neq i$ , or  $y_j^- \leq |x - y|^{1/2}$  for all  $j$  satisfying  $u_{ij} > \theta|x - y|^{1/4}$ . But the second situation does not occur due to (39). Thus for  $i \in \mathcal{A}^c(y)$  we have  $u_{ij} \leq \theta|x - y|^{1/4}$  for all  $j \neq i$ . Therefore there exists a constant  $\kappa_1$ , depending on  $\mathcal{H}$  and  $\theta$ , such that for all  $i \in \mathcal{A}^c(y)$ ,

$$\sum_{j: j \neq i} u_{ij} y_j^- \leq \kappa_1 |x - y|^{1/4}. \quad (40)$$

Now we define  $\bar{u} \in \mathbf{M}$  as follows

$$\bar{u}_{ij} = \begin{cases} 0, & \text{if } i \in \mathcal{A}^c(y), \\ 0, & \text{if } i \in \mathcal{A}(y) \text{ and } u_{ij} \leq \theta|x - y|^{1/4}, \\ (u_{ij} - \theta|x - y|^{1/4}), & \text{otherwise.} \end{cases}$$

We show that  $\bar{u} \in \mathbf{M}(x)$ . To do so we only need to check that for  $i \in \mathcal{A}(y)$ ,

$$\sum_{j: j \neq i} \bar{u}_{ij} x_j^- \leq x_i^+.$$

For  $i \in \mathcal{A}(y)$ ,

$$\begin{aligned} \sum_{j: j \neq i} \bar{u}_{ij} x_j^- &= \sum_{j: u_{ij} > \theta|x - y|^{1/4}} (u_{ij} - \theta|x - y|^{1/4}) x_j^- \\ &\leq \sum_{j: u_{ij} > \theta|x - y|^{1/4}} (u_{ij} - \theta|x - y|^{1/4}) (y_j^- + |x - y|) \\ &\leq \sum_{j: j \neq i} u_{ij} y_j^- + \sum_{j: u_{ij} > \theta|x - y|^{1/4}} (u_{ij} |x - y| - \theta|x - y|^{1/4} y_j^-) \\ &\leq y_i^+ + \sum_{j: u_{ij} > \theta|x - y|^{1/4}} (u_{ij} |x - y| - \theta|x - y|^{1/4} y_j^-) \\ &\leq x_i^+ + |x - y| \left( 1 + \sum_{j: j \neq i} u_{ij} \right) - \theta|x - y|^{3/4} \\ &\leq x_i^+, \end{aligned}$$

where in the last line we use the fact that  $|x - y|^{1/4} d \leq d/\theta^{1/4} < \theta$ . This shows  $\bar{u} \in \mathbf{M}(x)$ . Also by the construction of  $\bar{u}$  and (40) it is evident that for all  $i, j \in \{1, \dots, d\}$ ,  $i \neq j$ ,

$$|u_{ij} y_j^- - \bar{u}_{ij} x_j^-| \leq \kappa_2 |x - y|^{1/4}, \quad \text{and} \quad ||u_{ij}|^2 - |\bar{u}_{ij}|^2| \leq \kappa_2 |x - y|^{1/4},$$

for some constant  $\kappa_2$ . Hence (37) follows using the above display and (38).  $\square$

Now we are ready to prove Theorem 4.2.

**Proof of Theorem 4.2.** The proof is in spirit of Arapostathis et al. (2012, Theorem 3.5.6). Two key ingredients for the proof are Lemmas 4.1 and 4.3. Recall that  $\mathbb{B}_R$  denotes the ball of radius  $R$  around 0. It is known that there exists a solution  $\varphi_R \in \mathcal{C}^{2,\beta}(\mathbb{B}_R)$ ,  $\beta \in (0, 1/4)$ , satisfying

$$\begin{aligned} \frac{1}{2} \sum_{i,j=1}^d a_{ij} \partial_{ij} \varphi_R(x) + H_\varepsilon(x, \nabla \varphi(x)) &= \alpha \varphi_R(x), \quad x \in \mathbb{B}_R, \\ \varphi &= 0, \quad \text{on } x \in \partial \mathbb{B}_R. \end{aligned}$$

The existence result follows from Gilbarg and Trudinger (1983, Theorem 11.4) and maximum-principle together with Lemmas 4.1 and 4.3. See Arapostathis et al. (2012, Theorem 3.5.3) for a similar argument. Following a similar argument as in Arapostathis et al. (2012, Theorem 3.5.6) together with Lemma 4.1 we can show that

$$\varphi_R(x) = \inf_{U \in \Pi} \mathbb{E}_x \left[ \int_0^{\tau_R} e^{-\alpha s} \tilde{r}_\varepsilon(X(s), U(s)) ds \right]. \quad (41)$$

Therefore choosing  $k > m$  in Lemma 4.2 and using (13) we see that

$$V_\varepsilon^\alpha(x) \leq \inf_{U \in \Pi} \mathbb{E}_x \left[ \int_0^\infty e^{-\alpha s} h(X(s)) ds \right] < \frac{\kappa_1}{\alpha} + \mathcal{V}(x), \quad (42)$$

for some constant  $\kappa_1$ . Hence using (41) we get that  $\varphi_R(x) < \kappa_1/\alpha + \mathcal{V}(x)$  for all  $x$ . This shows that  $\varphi_R$  is locally bounded, uniformly in  $R$ . Hence using standard theory of elliptic PDE's and Lemma 4.1 we obtain that for any domain  $D \subset \mathbb{R}^d$  and  $p \geq 1$ ,

$$\|\varphi\|_{W^{2,p}(D)} \leq \kappa_2,$$

where the constant  $\kappa_2$  is independent of  $R$ . Hence we can extract a subsequence of  $\varphi_R$  that converges to  $\varphi$  as  $R \rightarrow \infty$  in  $\mathcal{W}_{\text{loc}}^{2,p}(\mathbb{R}^d)$ ,  $p \geq 2d$ , and in  $\mathcal{C}_{\text{loc}}^{1,r}(\mathbb{R}^d)$  for  $r \in (0, 1/4)$ . Using locally Lipschitz property of  $H_\varepsilon(x, \cdot)$ , by Lemma 4.3, we obtain that  $\varphi$  is a weak solution to

$$\min_{u \in M(x)} (L^u \varphi(x) + \tilde{r}_\varepsilon(x, u)) = \alpha \varphi(x), \quad \text{for } x \in \mathbb{R}^d. \quad (43)$$

Lemma 4.3 and the theory of elliptic PDE's give us  $\varphi \in \mathcal{C}^2(\mathbb{R}^d)$ . From (41) we also have  $\varphi \leq V_\varepsilon^\alpha$ . To show that equality we consider the minimizing selector  $u(\cdot)$  of (43). Existence of such selector is assured from Lemma 4.1. Hence we have

$$\frac{1}{2} \sum_{i,j=1}^d a_{ij} \partial_{i,j} \varphi + b(x, u(x)) \cdot \nabla \varphi(x) + \tilde{r}_\varepsilon(x, u(x)) = \alpha \varphi(x).$$

Since  $u(x) \in \mathbf{M}(x)$  for all  $x$ , the solution  $X$  corresponding to this Markov control is admissible. Therefore applying Itô's formula we get

$$\mathbb{E}_x[e^{-\alpha t \wedge \tau_R} \varphi(X(t \wedge \tau_R))] - \varphi(x) = \mathbb{E}_x^u \left[ \int_0^{t \wedge \tau_R} e^{-\alpha s} \tilde{r}_\varepsilon(X(s), u(X(s))) ds \right].$$

Now we use the non-negativity of  $\varphi$  to conclude that

$$\mathbb{E}_x^u \left[ \int_0^{t \wedge \tau_R} e^{-\alpha s} \tilde{r}_\varepsilon(X(s), u(X(s))) ds \right] \leq \varphi(x),$$

Letting  $R \rightarrow \infty$ , in the above display and using Fatou's lemma we get

$$\mathbb{E}_x^u \left[ \int_0^t e^{-\alpha s} \tilde{r}_\varepsilon(X(s), u(X(s))) ds \right] \leq \varphi(x).$$

This shows  $V_\varepsilon^\alpha = \varphi$  and  $u$  is an optimal Markov control.  $\square$

By  $\tilde{\mathbf{U}}_{\text{SM}}$  we denote the set of all admissible deterministic Markov controls, i.e.,  $\tilde{\mathbf{U}}_{\text{SM}}$  denotes collection of all measurable  $v: \mathbb{R}^d \rightarrow \mathbf{M}$  such that  $v(x) \in \mathbf{M}(x)$  for all  $x$ . We recall that  $\tilde{\tau}_r$  denotes the hitting time to the ball  $\mathbb{B}_r$ , i.e.,

$$\tilde{\tau}_r := \inf\{t \geq 0: X(t) \in \mathbb{B}_r\}.$$

**Lemma 4.4.** Let  $\bar{V}_\varepsilon^\alpha := V_\varepsilon^\alpha - V_\varepsilon^\alpha(0)$ . Then  $\bar{V}_\varepsilon^\alpha$  is bounded in  $\mathcal{W}_{\text{loc}}^{2,p}(\mathbb{R}^d)$ ,  $p \geq 1$ , and  $\{\alpha V_\varepsilon^\alpha(0)\}_{\alpha \in (0,1)}$  is also bounded. Let  $(V_\varepsilon, \varrho) \in \mathcal{W}_{\text{loc}}^{2,p}(\mathbb{R}^d) \times \mathbb{R}$  be any sub-sequential limit of  $(\bar{V}_\varepsilon^\alpha, \alpha V_\varepsilon^\alpha(0))$ , as  $\alpha \rightarrow 0$ , then we have  $V_\varepsilon \in \mathcal{C}^2(\mathbb{R}^d)$  that satisfies,

$$\min_{u \in M(x)} (L^u V_\varepsilon(x) + \tilde{r}_\varepsilon(x, u)) = \varrho. \quad (44)$$

Moreover  $\varrho \leq \varrho_\varepsilon$ . Furthermore,

- (a)  $V_\varepsilon(x) \leq \liminf_{r \downarrow 0} \inf_{v \in \tilde{\mathbf{U}}_{\text{SM}}} \mathbb{E}_x^v [\int_0^{\tilde{\tau}_r} (\tilde{r}_\varepsilon(X(s), v(X(s))) - \varrho) ds]$ .
- (b) If  $u_\varepsilon \in \tilde{\mathbf{U}}_{\text{SM}}$  denote the minimizing selector of (44) then

$$V_\varepsilon(x) \geq -\varrho \mathbb{E}_x^{u_\varepsilon} [\tilde{\tau}_r] - \sup_{\mathbb{B}_r} V_\varepsilon, \quad \text{for all } x \in \mathbb{B}_r^c.$$

**Proof.** From (42) we obtain that for any  $R > 0$ ,

$$\alpha \max_{x \in \mathbb{B}_R} V_\varepsilon^\alpha(x) < \kappa_R,$$

where the constant  $\kappa_R$  does not depend on  $\alpha$  and  $\varepsilon$ . Hence applying Arapostathis et al. (2012, Lemma 3.6.3) (see also Arapostathis et al. 2015, Lemma 3.5) we get that  $\bar{V}_\varepsilon^\alpha$  is bounded in  $\mathcal{W}_{\text{loc}}^{2,p}(\mathbb{R}^d)$ ,  $p \geq 1$ . Also boundedness  $\{\alpha V_\varepsilon^\alpha(0)\}_{\alpha \in (0,1)}$  follows from the above display. For  $p > 2d$ , we see that any sub-sequential limit also converges

in  $\mathcal{C}_{\text{loc}}^{1,r}(\mathbb{R}^d)$ . Hence any sub-sequential limit  $(V_\varepsilon, \varrho)$  would satisfy (44). We can improve regularity of  $V_\varepsilon$  to  $\mathcal{C}^2$  using Lemma 4.3. Now we show that  $\varrho \leq \varrho_\varepsilon$ . Let  $U \in \mathbb{U}$  be any admissible control. We also assume that

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_x^U \left[ \int_0^T \tilde{r}_\varepsilon(X(s), U(s)) ds \right] < \infty,$$

for some  $x$ . It is easy to see that

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_x^U \left[ \int_0^T \tilde{r}_\varepsilon(X(s), U(s)) ds \right] = \limsup_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}_x^U \left[ \int_0^N \tilde{r}_\varepsilon(X(s), U(s)) ds \right],$$

where  $N$  runs over natural numbers. In fact, it is easy to see that for RHS of above display is smaller than LHS, and for any  $T_n \rightarrow \infty$ , we have

$$\begin{aligned} \limsup_{T_n \rightarrow \infty} \frac{1}{T_n} \mathbb{E}_x^U \left[ \int_0^{T_n} \tilde{r}_\varepsilon(X(s), U(s)) ds \right] &= \limsup_{T_n \rightarrow \infty} \frac{\lfloor T_n + 1 \rfloor}{\lfloor T_n \rfloor} \frac{1}{\lfloor T_n + 1 \rfloor} \mathbb{E}_x^U \left[ \int_0^{\lfloor T_n + 1 \rfloor} \tilde{r}_\varepsilon(X(s), U(s)) ds \right] \\ &\leq \limsup_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}_x^U \left[ \int_0^N \tilde{r}_\varepsilon(X(s), U(s)) ds \right]. \end{aligned}$$

Define

$$a_N := \mathbb{E}_x^U \left[ \int_{N-1}^N \tilde{r}_\varepsilon(X(s), U(s)) ds \right], \quad N \geq 1.$$

For  $\beta = e^{-\alpha}$ , we have

$$\sum_{N=1}^{\infty} \beta^N a_N \geq e^{-\alpha} \sum_{N=1}^{\infty} \mathbb{E}_x^U \left[ \int_{N-1}^N e^{-\alpha s} \tilde{r}_\varepsilon(X(s), U(s)) ds \right] = e^{-\alpha} \mathbb{E}_x^U \left[ \int_0^{\infty} e^{-\alpha s} \tilde{r}_\varepsilon(X(s), U(s)) ds \right].$$

Therefore applying Sznajder and Filar (1992, Theorem 2.2) we get

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_x^U \left[ \int_0^T \tilde{r}_\varepsilon(X(s), U(s)) ds \right] \geq \limsup_{\alpha \rightarrow 0} (1 - \beta) \sum_{N=1}^{\infty} \beta^N a_N \geq \limsup_{\alpha \rightarrow 0} (1 - e^{-\alpha}) e^{-\alpha} V_\alpha^\varepsilon(x). \quad (45)$$

Since  $\lim_{\alpha \rightarrow 0} (1 - e^{-\alpha})/\alpha = 1$  we obtain from (45) that

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_x^U \left[ \int_0^T \tilde{r}_\varepsilon(X(s), U(s)) ds \right] \geq \varrho,$$

where we use the fact that for any  $x \in \mathbb{R}^d$ ,  $\alpha(V_\alpha^\varepsilon(x) - V_\varepsilon^\alpha(0)) \rightarrow 0$  as  $\alpha \rightarrow 0$ . Since  $U \in \mathbb{U}$  is arbitrary we have  $\varrho_\varepsilon \geq \varrho$ . (a) follows from the same argument as in Arapostathis et al. (2012, Lemma 3.7.8). In fact, following Arapostathis et al. (2012, Lemma 3.7.8) one obtains that for any  $r > 0$ ,

$$V_\varepsilon(x) \leq \inf_{v \in \mathbb{U}_{\text{SM}}} \mathbb{E}_x^v \left[ \int_0^{\tilde{\tau}_r} (\tilde{r}_\varepsilon(X(s), v(X(s))) - \varrho) ds + V_\varepsilon(X(\tilde{\tau}_r)) \right]. \quad (46)$$

Let  $u_\varepsilon^\alpha$  be the minimizing selector of (36). Then applying Lemma 4.1 we get that  $u_\varepsilon^\alpha \rightarrow u_\varepsilon$  pointwise, as  $\alpha \rightarrow 0$ . Also by Theorem 4.2 we have

$$\begin{aligned} \bar{V}_\varepsilon^\alpha(x) &= \mathbb{E}_x^{u_\varepsilon^\alpha} \left[ \int_0^{\tilde{\tau}_r} e^{-\alpha s} (\tilde{r}_\varepsilon(X(s), u_\varepsilon^\alpha(X(s))) - \varrho) ds \right] + \mathbb{E}_x^{u_\varepsilon^\alpha} [\bar{V}_\varepsilon^\alpha(X(\tilde{\tau}_r))] \\ &\quad + \mathbb{E}_x^{u_\varepsilon^\alpha} [\alpha^{-1} (1 - e^{-\alpha \tilde{\tau}_r}) (\varrho - \alpha V_\varepsilon^\alpha(X(\tilde{\tau}_r))], \quad \forall x \in \mathbb{B}_r^c. \end{aligned}$$

Since  $u_\varepsilon^\alpha \rightarrow u_\varepsilon$ , using Lemma 4.2 we have (see also Arapostathis et al. 2015, Lemma 3.8)

$$\mathbb{E}_x^{u_\varepsilon^\alpha} [\tilde{\tau}_r] \xrightarrow{\alpha \rightarrow 0} \mathbb{E}_x^{u_\varepsilon} [\tilde{\tau}_r], \quad \text{for all } x \in \mathbb{B}_r^c.$$

Since  $(1 - e^{-\alpha s}) \leq \alpha s$ , for  $s \geq 0$ , combining above two display we get, as  $\alpha \rightarrow 0$ , that

$$V_\varepsilon(x) \geq -\varrho \mathbb{E}_x^{u_\varepsilon} [\tilde{\tau}_r] - \sup_{y \in \mathbb{B}_r} V_\varepsilon(y), \quad \text{for } x \in \mathbb{B}_r^c. \quad \square$$

The following lemma establishes optimality of  $u_\varepsilon$  that we choose above.

**Lemma 4.5.** *Every sub-sequential limit  $V_\varepsilon$  that we obtain in Lemma 4.4 is in  $\mathcal{O}(|x|^k)$  for  $k > m$ , where  $m$  is given by (13). Also if  $u_\varepsilon$  is the minimizing selector in (44) we have*

$$\varrho_\varepsilon = \varrho = \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_x \left[ \int_0^T \tilde{r}_\varepsilon(X(s), u_\varepsilon(X(s))) ds \right], \quad \forall x \in \mathbb{R}^d. \quad (47)$$

**Proof.** From Lemma 4.4 (see (46)) we obtain that

$$|V_\varepsilon(x)| \leq \sup_{v \in \tilde{\mathcal{U}}_{SM}} \mathbb{E}_x^v \left[ \int_0^{\tilde{\tau}_r} (\tilde{r}_\varepsilon(X(s), v(X(s))) + \varrho_\varepsilon) ds \right] + \sup_{B_r} V_\varepsilon, \quad \forall x \in \mathbb{B}_r^c. \quad (48)$$

For  $v \in \tilde{\mathcal{U}}_{SM}$ , we define

$$L^v f(x) = \frac{1}{2} \sum_{i,j=1}^d a_{ij} \partial_{ij} f(x) + b(x, v(x)) \cdot \nabla f(x), \quad f \in \mathcal{C}^2(\mathbb{R}^d).$$

Considering  $k = m$  in Lemma 4.2 and applying Dynkin's formula we obtain, for  $v \in \tilde{\mathcal{U}}_{SM}$ , that

$$\begin{aligned} \mathbb{E}_x^v [\mathcal{V}(X(t \wedge \tilde{\tau}_r))] - \mathcal{V}(x) &= \mathbb{E}_x^v \left[ \int_0^{t \wedge \tilde{\tau}_r} L^v \mathcal{V}(X(s)) ds \right] \\ &= \mathbb{E}_x^v \left[ \int_0^{t \wedge \tilde{\tau}_r} 1_{\{v(X(s)) \in M(X(s))\}} L^v \mathcal{V}(X(s)) ds \right] \leq \mathbb{E}_x^v \left[ \int_0^{t \wedge \tilde{\tau}_r} (c_5 - c_6 |X(s)|^m) ds \right]. \end{aligned}$$

Now  $\mathcal{V}$  being non-negative, letting  $t \rightarrow \infty$ , in the above display we obtain

$$\mathbb{E}_x^v \left[ \int_0^{\tilde{\tau}_r} \left( |X(s)|^m - \frac{c_5}{c_6} \right) ds \right] \leq \frac{1}{c_6} \mathcal{V}(x), \quad \text{for } x \in \mathbb{B}_r^c.$$

Choose  $r \geq$  large enough so that  $|x| \geq \frac{1}{2} \vee 4c_5/c_6$  for  $x \in \mathbb{B}_r^c$ . Therefore

$$\mathbb{E}_x^v \left[ \int_0^{\tilde{\tau}_r} \frac{1}{4} (1 + |X(s)|^m) ds \right] \leq \mathbb{E}_x^v \left[ \int_0^{\tilde{\tau}_r} \frac{1}{2} |X(s)|^m ds \right] \leq \frac{1}{c_6} \mathcal{V}(x), \quad \text{for } x \in \mathbb{B}_r^c.$$

Thus (13) and (48) gives that  $V_\varepsilon \in \mathcal{O}(|x|^m)$ . Hence  $V_\varepsilon \in \mathcal{O}(|x|^k)$ ,  $k > m$ .

Now let  $u_\varepsilon$  be the minimizing selector of (44). We observe that  $u_\varepsilon \in \tilde{\mathcal{U}}_{SM}$ . Moreover, using Lemma 4.2 we see that  $u_\varepsilon$  is stable with

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_x^{u_\varepsilon} \left[ \int_0^T |X(t)|^k dt \right] < \infty,$$

for any  $k \geq 1$ . Thus if  $\mu_{u_\varepsilon}$  denote the invariant measure corresponding to the Markov control  $u_\varepsilon$  we have

$$\int_{\mathbb{R}^d} |x|^k \mu_{u_\varepsilon}(dx) < \infty, \quad \text{for } k \geq 1.$$

Since  $V_\varepsilon \in \mathcal{O}(|x|^k)$ ,  $k > m$ , it follows from Ichihara and Sheu (2013, Proposition 2.6) that

$$\lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_x^{u_\varepsilon} [|V_\varepsilon(X(T))|] = 0, \quad \text{for } x \in \mathbb{R}^d.$$

Thus (47) follows by an application of Dynkin's formula to (44).  $\square$

To this end we define

$$H(p, x) := \inf_{u \in M(x)} \{p \cdot b(x, u) + \tilde{r}(x, u)\}.$$

the following result is similar to Lemma 4.3.

**Lemma 4.6.** *Let  $\mathcal{K} \subset \mathbb{R}^d$  be compact. Then for any  $R > 0$ , there exists a constant  $c = c_{\mathcal{K}, R}$ , depending on  $\mathcal{K}$ ,  $R$ , such that*

$$|H(x, p) - H(y, q)| \leq c(|x - y|^{1/2} + |p - q|), \quad \text{for all } x, y \in \mathcal{K}, \text{ and, } p, q \in \mathbb{B}_R(0).$$

**Proof.** We note that the proof of Lemma 4.3 is not applicable here to conclude the result. In fact, the constant  $\theta$  defined in Lemma 4.3 tends to infinity as  $\varepsilon \rightarrow 0$ . However, we adopt a similar technique to establish Hölder regularity of the Hamiltonian. In view of Lemma 4.3, we see that it is enough to show the following: there exists  $\theta_1$ , depending on  $\mathcal{H}, R$ , such that for any  $u \in \mathbf{M}(y)$  we could find  $\bar{u} \in \mathbf{M}(x)$  such that

$$|u_{ij}y_j^- - \bar{u}_{ij}x_j^-| \leq \theta_1|x-y|^{1/2}, \quad \text{for all } i \neq j. \quad (49)$$

and  $x, y \in \mathcal{H}$ . Let  $u \in \mathbf{M}(y)$ . Let  $\theta$  be a positive number that will be chosen later. Define

$$\mathcal{A}(\theta) := \left\{ i \in \{1, \dots, d\} : \sum_{j: u_{ij} > \theta|x-y|^{1/2}} y_j^- \leq |x-y|^{1/2} \right\}.$$

We define  $\bar{u} \in \mathbf{M}$  as follows,

$$\bar{u}_{ij} = \begin{cases} 0, & \text{if } i \in \mathcal{A}(\theta), \\ 0, & \text{if } i \in \mathcal{A}^c(\theta), \text{ and } u_{ij} \leq \theta|x-y|^{1/2}, \\ u_{ij} - \theta|x-y|^{1/2}, & \text{otherwise.} \end{cases}$$

It is easy to see that (49) is satisfied with the above choice of  $\bar{u}$ . Thus it remains to show that  $\bar{u} \in \mathbf{M}(x)$ . It is enough to show that for  $i \in \mathcal{A}^c(\theta)$ , we have

$$\sum_{j: j \neq i} \bar{u}_{ij}x_j^- \leq x_i^+. \quad (50)$$

Now for  $i \in \mathcal{A}^c(\theta)$ ,

$$\begin{aligned} \sum_{j: j \neq i} \bar{u}_{ij}x_j^- &\leq \sum_{u_{ij} > \theta|x-y|^{1/2}} (u_{ij} - \theta|x-y|^{1/2})(y_j^- + |x-y|) \\ &\leq \sum_{u_{ij} > \theta|x-y|^{1/2}} u_{ij}y_j^- + |x-y| \sum_{j: j \neq i} u_{ij} - \theta|x-y|^{1/2} \sum_{u_{ij} > \theta|x-y|^{1/2}} y_j^- \\ &\leq y_i^+ + |x-y| \sum_{j: j \neq i} u_{ij} - \theta|x-y|^{1/2} \sum_{u_{ij} > \theta|x-y|^{1/2}} y_j^- \\ &\leq x_i^+ + d|x-y| - \theta|x-y|. \end{aligned}$$

Thus choosing  $\theta > d$  we see that (50) holds for  $i \in \mathcal{A}^c(\theta)$  and  $\bar{u} \in \mathbf{M}(x)$ .  $\square$

Now we are ready to prove Theorem 3.1.

**Proof of Theorem 3.1.** Existence of solution  $V_\varepsilon \in \mathcal{C}^2$  and optimality of  $\rho_\varepsilon$  follows from Lemma 4.5. Uniqueness of  $(V_\varepsilon, \rho_\varepsilon)$  can be obtained following the same arguments as in Arapostathis et al. (2012, Theorem 3.7.12(iii)). Hence (i) and (ii) follows. Now we argue that for any  $r > 0$ ,

$$\sup_{\varepsilon \in (0,1)} \sup_{\mathbb{B}_r} |V_\varepsilon| < \infty. \quad (51)$$

We recall that  $V_\varepsilon$  is obtained as a limit of  $\bar{V}_\varepsilon^\alpha = V_\varepsilon^\alpha - V_\varepsilon^\alpha(0)$  where  $V_\varepsilon^\alpha$  is given by (35). From Arapostathis et al. (2015, Lemma 3.5) (see also Arapostathis et al. 2012, Lemma 3.6.3) one obtains that  $\sup_{\varepsilon \in (0,1)} \sup_{x \in \mathbb{B}_r} |\bar{V}_\varepsilon^\alpha(x)| < \infty$ . This gives us (51). Therefore combining (51), (48) and the arguments in Lemma 4.5 we find a constant  $\kappa$  such that

$$\sup_{\varepsilon \in (0,1)} |V_\varepsilon(x)| \leq \kappa(1 + |x|^m), \quad \forall x \in \mathbb{R}^d.$$

Hence applying the theory of elliptic PDE's we see that the family  $\{V_\varepsilon\}$  bounded in  $\mathcal{W}_{\text{loc}}^{2,p}(\mathbb{R}^d)$ , for  $p \geq d$ . Since  $\mathcal{W}_{\text{loc}}^{2,p}(\mathbb{R}^d)$ ,  $p > 2d$ , is compactly embedded in  $\mathcal{C}_{\text{loc}}^{1,\beta}(\mathbb{R}^d)$ ,  $\beta \in (0, 1/2)$ , we obtain  $\{V_\varepsilon\}$  bounded in  $\mathcal{C}_{\text{loc}}^{1,\beta}(\mathbb{R}^d)$ ,  $\beta \in (0, 1/2)$ . Thus we have  $V \in \mathcal{W}_{\text{loc}}^{2,p}(\mathbb{R}^d) \cap \mathcal{C}_{\text{loc}}^{1,\beta}(\mathbb{R}^d)$  with  $p > 2d$ , and  $\beta \in (0, 1/2)$ , such that  $V_\varepsilon \rightarrow V$  in  $\mathcal{W}_{\text{loc}}^{2,p}(\mathbb{R}^d) \cap \mathcal{C}_{\text{loc}}^{1,\beta}(\mathbb{R}^d)$  along some sub-sequence of  $\varepsilon \rightarrow 0$ . Letting  $\varepsilon \rightarrow 0$  in (28) we see that  $(V, \rho)$  satisfies (29). Using Lemma 4.6 and regularity property of non-degenerate elliptic operator we get  $V \in \mathcal{C}^2(\mathbb{R}^d)$ .

(iv): Therefore there exists a classical solution  $(V, \rho_*)$  to (29). We now show that there exists a measurable minimizing selector of (29). We first show that the map  $\chi: \mathbb{R}^d \rightarrow 2^M$ , defined as  $\chi(x) = \mathbf{M}(x)$ , is measurable. To check measurability we need to show that for any closed  $F \subset M$ ,

$$\chi^\ell(F) := \{x \in \mathbb{R}^d : \chi(x) \cap F \neq \emptyset\},$$

is a Borel set (see Aliprantis and Border 2006, p. 557).  $\chi^\ell$  is referred to as the *lower-inverse* of  $\chi$ . In fact, we show that  $\chi^\ell(F)$  is a closed set whenever  $F$  is closed. Let  $x_n \rightarrow x$ , as  $n \rightarrow \infty$ , for a sequence  $\{x_n\} \subset \chi^\ell(F)$ . Then there exists  $u_n$  such that  $u_n \in \chi(x_n) \cap F$  and  $u_n \in \mathbf{M}(x_n)$ . Now  $\mathbf{M}$  being compact there exists  $u \in F$  satisfying  $u_n \rightarrow u$  as  $n \rightarrow \infty$ . It is easy to check that  $u \in \mathbf{M}(x)$ . Thus  $\chi(x) \cap F \neq \emptyset$  implying  $x \in \chi^\ell(F)$ . This shows that  $\chi^\ell(F)$  is a Borel set. Thus  $\chi$  is a measurable correspondence (Aliprantis and Border 2006, Chapter 18). Since  $\chi$  is a compact set-valued map it is a weakly-measurable correspondence (Aliprantis and Border 2006, Theorem 18.10). Hence from Filippov's implicit function theorem (Aliprantis and Border 2006, Theorem 18.17) we obtain that there exists a measurable selector  $u: \mathbb{R}^d \rightarrow \mathbf{M}$  such that  $u(x) \in \mathbf{M}(x)$  for all  $x$ , and

$$\frac{1}{2} \sum_{i,j=1}^d a_{ij} \partial_{ij} V(x) + b(x, u(x)) \cdot \nabla V(x) + \tilde{r}(x, u(x)) = \varrho_*. \quad (52)$$

On the other hand, we have  $V \in \mathfrak{v}(|x|^k)$ ,  $k > m$ . Therefore applying Dynkin's formula to (52) with a similar argument as in Lemma 4.5 we obtain that  $u$  is optimal for  $\varrho_*$ .

(iii) and (v): We have already shown above that  $V_\varepsilon \rightarrow V$ , along some subsequence as  $\varepsilon \rightarrow 0$ , in  $\mathcal{W}_{\text{loc}}^{2,p}(\mathbb{R}^d)$ ,  $p \geq 1$ . Therefore to get the convergence of full sequence it is enough to establish the uniqueness of the limit. From Lemma 4.4(a) and (46) we observe that

$$V(x) \leq \liminf_{r \downarrow 0} \inf_{v \in \tilde{\mathbf{U}}_{\text{SM}}} \mathbb{E}_x^v \left[ \int_0^{\tilde{\tau}_r} (\tilde{r}(X(s), v(X(s))) - \varrho_*) ds \right].$$

Now uniqueness can be obtained following a similar argument as in Arapostathis et al. (2012, Theorem 3.7.12(iii)).  $\square$

## 5. Asymptotic Optimality

The goal of this section is to prove Theorem 2.1. In Theorem 5.1 below we show that  $\varrho_*$  is an asymptotic lower bound for the value functions  $\hat{V}^n(\hat{X}^n(0))$  as  $n \rightarrow \infty$ . Then using Theorem 3.1 we construct a sequence of admissible policies for the queueing systems and show in Theorem 5.2 that the admissible policies are  $\varepsilon$ -optimal for the value functions as  $n \rightarrow \infty$ .

Recall the diffusion scaled process  $\hat{X}^n$ ,  $\hat{Z}^n$  and  $\hat{Q}^n$  from (7) and their relation (8)

$$\begin{aligned} \hat{X}_i^n(t) &= \hat{X}_i^n(0) + \ell_i^n t + \mu_i^n \int_0^t (\hat{X}_i^n)^-(s) ds - \sum_{j:j \neq i} \mu_{ij}^n \int_0^t \hat{Z}_{ij}^n(s) ds - \gamma_i^n \int_0^t \hat{Q}_i^n(s) ds \\ &\quad + \hat{M}_{A,i}^n(t) - \hat{M}_{S,i}^n(t) - \sum_{j:j \neq i} \hat{M}_{S,ij}^n(t) - \hat{M}_{R,i}^n(t), \end{aligned} \quad (53)$$

where  $\ell^n = (\ell_1^n, \dots, \ell_d^n)^\top$  is defined as

$$\ell_i^n := \frac{1}{\sqrt{n}} (\lambda_i^n - \mu_i^n n) = \frac{\lambda_i^n - n \lambda_i}{\sqrt{n}} - \sqrt{n} (\mu_i^n - \mu_i),$$

and (see (9))

$$\begin{aligned} \hat{M}_{A,i}^n(t) &:= \frac{1}{\sqrt{n}} (A_i^n(\lambda_i^n t) - \lambda_i^n t), & \hat{M}_{S,i}^n(t) &:= \frac{1}{\sqrt{n}} \left( S_i^n \left( \mu_i^n \int_0^t Z_i^n(s) ds \right) - \mu_i^n \int_0^t Z_i^n(s) ds \right), \\ \hat{M}_{S,ij}^n(t) &:= \frac{1}{\sqrt{n}} \left( S_{ij}^n \left( \mu_{ij}^n \int_0^t Z_{ij}^n(s) ds \right) - \mu_{ij}^n \int_0^t Z_{ij}^n(s) ds \right), & \hat{M}_{R,i}^n(t) &:= \frac{1}{\sqrt{n}} \left( R_i^n \left( \gamma_i^n \int_0^t Q_i^n(s) ds \right) - \gamma_i^n \int_0^t Q_i^n(s) ds \right), \end{aligned}$$

are square integrable martingales w.r.t. the filtration  $\{\mathcal{F}_t^n\}$  with quadratic variations,

$$\langle \hat{M}_{A,i}^n \rangle(t) := \frac{1}{n} \lambda_i^n t, \quad \langle \hat{M}_{S,i}^n \rangle(t) := \frac{1}{n} \mu_i^n \int_0^t Z_i^n(s) ds, \quad \langle \hat{M}_{S,ij}^n \rangle(t) := \frac{1}{n} \mu_{ij}^n \int_0^t Z_{ij}^n(s) ds, \quad \langle \hat{M}_{R,i}^n \rangle(t) := \frac{1}{n} \gamma_i^n \int_0^t Q_i^n(s) ds.$$

Also recall  $\hat{U}^n(t)$  from (12) where we have  $\hat{Z}_{ij}^n = \hat{U}_{ij}^n(\hat{X}_j^n)^-$ . We also have  $\hat{U}(t) \in \mathbf{M}(\hat{X}^n(t))$  a.s. for all  $t \geq 0$ . Define  $b^n: \mathbb{R}^d \times \mathbf{M} \rightarrow \mathbb{R}^d$  as

$$b_i^n(x, u) = \ell_i^n + \mu_i^n x_i^- - \sum_{j:j \neq i} \mu_{ij}^n u_{ij} x_j^- - \gamma_i^n q(x, u),$$

where

$$q_i(x, u) = x_i^+ - \sum_{j:j \neq i} u_{ij} x_j^-.$$

By  $\mathbb{1}^n$  we denote the set of all admissible controls. The following result establishes uniform stability of the queueing systems.

**Lemma 5.1.** *There exists  $n_0 \geq 1$ , such that for any  $k \geq 1$ ,*

$$\sup_{n \geq n_0} \sup_{Z \in \mathbb{1}^n} \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[ \int_0^T |\hat{X}^n(t)|^k dt \right] < \infty. \quad (54)$$

**Proof.** Without loss of generality, we take  $k \geq 2$ . Let  $\varphi(x) := |x|^k$ . Recall that  $\Delta f(t)$  denotes jump of a function  $f: \mathbb{R}_+ \rightarrow \mathbb{R}$  at time  $t$ . Components of  $\hat{X}^n$  jumps due to the jumps of their martingale parts. Since the optional quadratic variation between martingales that corresponds to different components is 0 (see Pang et al. 2007, Lemma 9.1) no two component jumps at the same time. Now applying Itô's formula on  $\varphi$  (see Kallenberg 2002, Theorem 26.7) we obtain from (53) that

$$\begin{aligned} \mathbb{E}[\varphi(\hat{X}^n(t))] &= \mathbb{E}[\varphi(\hat{X}^n(0))] + \mathbb{E} \left[ \int_0^t b^n(\hat{X}^n(s), \hat{U}^n(s)) \cdot \nabla \varphi(\hat{X}^n(s)) ds \right] \\ &\quad + \frac{1}{2} \sum_{i=1}^d \mathbb{E} \left[ \int_0^t \Theta_i(\hat{Z}^n(s), \hat{Q}^n) \cdot \partial_{ii} \varphi(\hat{X}^n(s)) ds \right] \\ &\quad + \mathbb{E} \sum_{s \leq t} \left[ \Delta \varphi(\hat{X}(s)) - \sum_{i=1}^d \partial_i \varphi(\hat{X}(s-)) \Delta \hat{X}_i^n(s) - \frac{1}{2} \langle \Delta \hat{X}^n(s), D^2 \varphi(\hat{X}^n(s-)) \Delta \hat{X}^n(s) \rangle \right], \end{aligned} \quad (55)$$

where  $D^2 \varphi$  denotes Hessian of  $\varphi$  and,

$$\Theta_i(z, y) := \frac{1}{n} \lambda_i^n + \frac{1}{\sqrt{n}} \mu_i^n z_i^n + \mu_i^n + \frac{1}{\sqrt{n}} \sum_{j: j \neq i} z_{ij} + \frac{1}{\sqrt{n}} \gamma_i^n y_i.$$

Using (2) we can choose  $n$  large enough so that  $\min_i (\mu_i^N \wedge \gamma_i^n) > 0$ . An argument similar to (32)–(33) shows that for some positive constants  $\kappa_1, \kappa_2$ , independent of  $n$ , we have

$$b^n(\hat{X}^n(t), \hat{U}^n(t)) \cdot \nabla \varphi(\hat{X}^n(t)) \leq \kappa_1 - \kappa_2 |\hat{X}^n(t)|^k, \quad \text{a.s., for all } t \geq 0. \quad (56)$$

Also from (6) we obtain  $\hat{Z}_i^n = (\hat{X}_i^n)^-$  for all  $i$ . Thus for all  $t \geq 0$ , we have

$$|\Theta_i(\hat{Z}^n(t), \hat{Q}^n(t)) \cdot \partial_{ii} \varphi(\hat{X}^n(t))| \leq \kappa_3 (1 + |\hat{X}^n(t)|^{k-1}), \quad \text{for all } i \in \{1, \dots, d\}, \quad (57)$$

for some constant  $\kappa_3$  and all large  $n$ , where we use (5). We observe that  $|\Delta \hat{X}^n(t)| \leq 1/\sqrt{n}$  for all  $t$ . Hence a straightforward calculation gives us

$$\left[ \Delta \varphi(\hat{X}(s)) - \sum_{i=1}^d \partial_i \varphi(\hat{X}(s-)) \Delta \hat{X}_i^n(s) - \frac{1}{2} \langle \Delta \hat{X}^n(s), D^2 \varphi(\hat{X}^n(s-)) \Delta \hat{X}^n(s) \rangle \right] \leq \frac{\kappa_4}{\sqrt{n}} (1 + |\hat{X}^n(s-)|^{k-2}) \sum_{i=1}^d (\Delta \hat{X}_i^n(s))^2. \quad (58)$$

Since  $\sum_{s \leq t} (\Delta \hat{X}_i^n(s))^2 := [\hat{X}_i^n](t)$  where  $[\hat{X}_i^n]$  is the optional quadratic variation, we get from (58) that

$$\begin{aligned} &\mathbb{E} \sum_{s \leq t} \left[ \Delta \varphi(\hat{X}(s)) - \sum_{i=1}^d \partial_i \varphi(\hat{X}(s-)) \Delta \hat{X}_i^n(s) - \frac{1}{2} \langle \Delta \hat{X}^n(s), D^2 \varphi(\hat{X}^n(s-)) \Delta \hat{X}^n(s) \rangle \right] \\ &\leq \frac{\kappa_4}{\sqrt{n}} \sum_{i=1}^d \mathbb{E} \left[ \int_0^t \Theta_i(\hat{Z}^n(s), \hat{Q}^n) (1 + |\hat{X}_i^n(s)|^{k-2}) ds \right] \leq \frac{\kappa_5}{\sqrt{n}} \left( t + \mathbb{E} \left[ \int_0^t |\hat{X}^n(s)|^{k-1} ds \right] \right), \end{aligned} \quad (59)$$

for some constant  $\kappa_5$ , where we use the fact that  $[\hat{X}_i^n] - \langle \hat{X}_i^n \rangle$  is also a martingale. Therefore combining (55)–(57), and (59) we obtain constants  $\kappa_6, \kappa_7 > 0$ , independent of  $\mathbb{1}^n$ , such that

$$\mathbb{E}[\varphi(\hat{X}^n(t))] = \mathbb{E}[\varphi(\hat{X}^n(0))] + \kappa_6 t - \kappa_7 \mathbb{E} \left[ \int_0^t |\hat{X}^n(s)| ds \right],$$

for all large  $n$ . Since  $\{\hat{X}^n(0)\}$  is bounded, we obtain (54) from the above display.  $\square$

For any  $n \geq 1$  and  $(\hat{X}^n, \hat{U}^n)$  satisfying (8) and (12) we define mean-empirical measures  $\xi_t^n \in \mathcal{P}(\mathbb{R}^d \times \mathbf{M})$  as follows: For Borel  $A \in \mathcal{B}(\mathbb{R}^d)$ ,  $B \in \mathbf{M}$ ,

$$\xi_t^n(A \times B) := \frac{1}{t} \mathbb{E} \left[ \int_0^t 1_{A \times B}(\hat{X}^n(s), \hat{U}^n(s)) ds \right], \quad t > 0. \quad (60)$$

Let

$$\tilde{\Theta}_i^n := \frac{1}{n} \lambda_i^n + \frac{1}{\sqrt{n}} \mu_i^n x_i^- + \mu_i^n + \frac{1}{\sqrt{n}} \sum_{j:j \neq i} u_{ij} x_j^- + \frac{1}{\sqrt{n}} \gamma_i^n q_i(x, u),$$

where  $q$  is given by (20). From (2) we have  $\tilde{\Theta}_i^n \rightarrow 2\lambda_i$ , as  $n \rightarrow \infty$ , uniformly on compacts.

**Lemma 5.2.** Consider  $n \geq n_0$ , where  $n_0$  is given by Lemma 5.1. For  $\hat{U}^n \in \mathbb{U}^n$ , we define  $\xi_t^n$  as in (60). Then the collection  $\{\xi_t^n, t > 0\}$  is tight, as  $t \rightarrow \infty$ , and if  $\pi^n$  is a sub-sequential limit of  $\{\xi_t^n, t > 0\}$ , then we have

$$\int_{\mathbb{R}^d \times \mathbf{M}} 1_{M^c(x)}(u) \pi^n(dx, du) = 0.$$

**Proof.** From (54) we obtain that

$$\limsup_{T \rightarrow \infty} \int_{\mathbb{R}^d \times \mathbf{M}} |x|^k \xi_T^n(dx, du) < \infty,$$

for any  $k \geq 1$ . This implies that  $\{\xi_t^n, t > 0\}$  is tight. Let  $\pi^n$  be a sub-sequential limit and  $\xi_{t_l}^n \rightarrow \pi^n$  as  $t_l \rightarrow \infty$ . Since  $\hat{U}^n(\hat{X}^n(s)) \in \mathbf{M}(\hat{X}^n(s))$  for all  $s$ , we have from (60) that

$$\int_{\mathbb{R}^d \times \mathbf{M}} 1_{M^c(x)}(u) \xi_{t_l}^n(dx, du) = 0, \quad \forall t > 0. \quad (61)$$

We note that  $(x, u) \mapsto 1_{M^c(x)}(u)$  is a lower-semicontinuous function. Thus there exists a sequence of bounded, continuous functions  $g_j$  such that  $g_j(x, u) \nearrow 1_{M^c(x)}(u)$  as  $j \rightarrow \infty$  (Krantz 2001, Proposition 2.1.2). Therefore (61) gives

$$\int_{\mathbb{R}^d \times \mathbf{M}} g_j(x, u) \pi^n(dx, du) = \lim_{t_l \rightarrow \infty} \int_{\mathbb{R}^d \times \mathbf{M}} g_j(x, u) \xi_{t_l}^n(dx, du) = 0.$$

Now let  $j \rightarrow \infty$ , to complete the proof.  $\square$

Now we establish asymptotic lower-bound of the value functions  $\hat{V}^n$ .

**Theorem 5.1.** As  $n \rightarrow \infty$ ,  $\hat{V}^n(\hat{X}^n(0)) \geq \varrho_*$ , where  $\varrho_*$  is given by (27).

**Proof.** Consider a sequence  $Z^n \in \mathbb{U}^n$  and let  $\{\xi_t^n\}$  be the associated mean-empirical measures as defined in (60). From (5.2) we obtain that for  $n \geq n_0$ , the collection  $\{\xi_t^n\}$  is tight. Let  $\pi^n$  be a subsequential limit of  $\{\xi_t^n\}$  as  $t \rightarrow \infty$ . Taking  $k > m$  in Lemma 5.1, we obtain that

$$\sup_{n \geq n_0} \int_{\mathbb{R}^d \times \mathbf{M}} \tilde{r}(x, u) \pi^n(dx, du) < \infty, \quad \text{and,} \quad \sup_{n \geq n_0} \int_{\mathbb{R}^d \times \mathbf{M}} |x|^k \pi^n(dx, du) < \infty. \quad (62)$$

Thus from (62) we see that the sequence  $\{\pi^n: n \geq 1\}$  is also a tight sequence. Let  $\pi$  be a sub-sequential limit of  $\{\pi^n: n \geq 1\}$  as  $n \rightarrow \infty$ . We show that  $\pi \in \mathcal{G}$  where  $\mathcal{G}$  is given by (34). Consider  $f \in \mathcal{C}_c^2(\mathbb{R}^d)$  and apply Itô's formula in (53) to obtain

$$\begin{aligned} \mathbb{E}[f(\hat{X}^n(t))] &= \mathbb{E}[f(\hat{X}^n(0))] + \mathbb{E} \left[ \int_0^t b^n(\hat{X}^n(s), \hat{U}^n(s)) \cdot \nabla f(\hat{X}^n(s)) ds \right] \\ &\quad + \frac{1}{2} \sum_{i=1}^d \mathbb{E} \left[ \int_0^t \tilde{\Theta}_i^n(\hat{X}^n(s), \hat{U}^n(s)) \cdot \partial_{ii} f(\hat{X}^n(s)) ds \right] \\ &\quad + \mathbb{E} \sum_{s \leq t} \left[ \Delta f(\hat{X}(s)) - \sum_{i=1}^d \partial_i f(\hat{X}(s-)) \Delta \hat{X}_i^n(s) - \frac{1}{2} \langle \Delta \hat{X}^n(s), D^2 f(\hat{X}^n(s-)) \Delta \hat{X}^n(s) \rangle \right], \end{aligned}$$

and therefore dividing by  $t$ , we get

$$\begin{aligned} \frac{\mathbb{E}[f(\hat{X}^n(t))]}{t} &= \frac{\mathbb{E}[f(\hat{X}^n(0))]}{t} + \int_{\mathbb{R}^d \times M} b^n(x, u) \cdot \nabla f(x) \xi_t^n(dx, du) \\ &\quad + \frac{1}{2} \sum_{i=1}^d \int_{\mathbb{R}^d \times M} \tilde{\Theta}_i^n(x, u) \cdot \partial_{ii} f(x) \xi_t^n(dx, du) \\ &\quad \cdot \frac{1}{t} \mathbb{E} \sum_{s \leq t} \left[ \Delta f(\hat{X}(s)) - \sum_{i=1}^d \partial_i f(\hat{X}(s-)) \Delta \hat{X}_i^n(s) - \frac{1}{2} \langle \Delta \hat{X}^n(s), D^2 f(\hat{X}^n(s-)) \Delta \hat{X}^n(s) \rangle \right]. \end{aligned}$$

Now let  $t \rightarrow \infty$ , and use a similar argument as in (59) to obtain

$$\int_{\mathbb{R}^d \times M} \left( \frac{1}{2} \sum_{i=1}^d \tilde{\Theta}_i^n(x, u) \cdot \partial_{ii} f(x) + b^n(x, u) \cdot \nabla f(x) \right) \pi^n(dx, d(u)) = \mathcal{O}\left(\frac{1}{\sqrt{n}}\right). \quad (63)$$

therefore letting  $n \rightarrow \infty$ , in (63) and using locally uniform convergence property of  $\tilde{\Theta}^n, b^n$  we get

$$\int_{\mathbb{R}^d \times M} L^u f(x) \pi(dx, d(u)) = 0,$$

where  $L^u$  is given by (24). Therefore to show  $\pi \in \mathcal{G}$  it remains to prove that

$$\int_{\mathbb{R}^d \times M} 1_{M^c(x)}(u) \pi(dx, du) = 0.$$

But this follows using the second part of Lemma 5.2 and lower semicontinuity of the map. We also have

$$\liminf_{n \rightarrow \infty} \int_{\mathbb{R}^d \times M} \tilde{r}(x, u) \pi^n(dx, du) \geq \int_{\mathbb{R}^d \times M} \tilde{r}(x, u) \pi(dx, du).$$

Hence using Theorem 4.1, we conclude that

$$\liminf_{n \rightarrow \infty} \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[ \int_0^T r(\hat{Q}^n(s)) ds \right] \geq \varrho_*. \quad \square$$

Next we proceed to prove the asymptotic upper bound. The idea is to construct a sequence of admissible policies that achieves  $\varrho_*$ . One main obstacle with such construction is that the minimizer of the HJB (29) might not be continuous, in general. Therefore we use the perturbed HJB (28). Let  $u: \mathbb{R}^d \rightarrow \mathbf{M}$  be a continuous function and  $u(x) \in \mathbf{M}(x)$  for all  $x$ . Using  $u$  we construct an admissible policy for every  $n$  as follows. Recall that  $\lfloor a \rfloor$  denotes the largest integer small or equal to  $a \in \mathbb{R}$ . For  $X^n(t) \in \mathbb{R}_+^d$ , we define,

$$Z_i^n(t) := X^n(t) \wedge n, \quad Z_{ki}^n(t) := \lfloor u_{ki}(\hat{X}^n(t))(X_i^n(t) - n)^- \rfloor, \quad i \neq k,$$

where  $\hat{X}^n$  denotes the scaled version of  $X^n$  under diffusion settings. We also define

$$Q_i^n(t) = X_i^n(t) - Z_i^n(t) - \sum_{j: j \neq i} Z_{ij}^n(t), \quad \text{for } i \in \{1, \dots, d\}.$$

We check that for  $i \in \{1, \dots, d\}$

$$Z_i^n(t) + \sum_{k: k \neq i} Z_{ki}^n(t) \leq X^n(t) \wedge n + \sum_{k: k \neq i} u_{ki}(\hat{X}^n(t))(X_i^n(t) - n)^- \leq X^n(t) \wedge n + (X_i^n(t) - n)^- \leq n,$$

and

$$\sum_{j: j \neq i} Z_{ij}^n(t) \leq \sum_{j: j \neq i} u_{ij}(\hat{X}^n(t))(X_j^n(t) - n)^- = \sqrt{n} \sum_{j: j \neq i} u_{ij}(\hat{X}^n(t))(\hat{X}_j^n(t))^- \leq (X_i^n(t) - n)^+.$$

Therefore  $Z^n \in \mathfrak{U}$  for all  $n$ . It is easy to see that

$$|\hat{Z}_{ij}^n - u_{ij}(\hat{X}_j^n)(\hat{X}_j^n)^-| \leq \frac{1}{\sqrt{n}}, \quad \text{for all } i \neq j.$$

**Theorem 5.2.** *We have*

$$\limsup_{n \rightarrow \infty} \hat{V}^n(\hat{X}^n(0)) \leq \varrho_*,$$

where  $\varrho_*$  is given by (27).

**Proof.** Since  $\varrho_\varepsilon \searrow \varrho_*$ , as  $\varepsilon \rightarrow 0$ , it is enough to show that there exists a sequence of admissible policy  $Z^n \in \mathbb{U}^n$  satisfying

$$\limsup_{n \rightarrow \infty} J(\hat{X}^n(0), \hat{Z}^n) \leq \varrho_\varepsilon, \quad (64)$$

where  $J(\hat{X}^n(0), \hat{Z}^n)$  is defined in (14). Let  $u_\varepsilon$  be the minimizing selector of (28). Now we construct a sequence of policy  $Z^n$  as above given  $u_\varepsilon$ . We define empirical measures  $\tilde{\xi}_t^n \in \mathcal{P}(\mathbb{R}^d)$  as

$$\tilde{\xi}_t^n(A) := \frac{1}{t} \mathbb{E}_x \left[ \int_0^t 1_A(\hat{X}^n(s)) ds \right], \quad t > 0.$$

From Lemma 5.1 we see that  $\{\tilde{\xi}_t^n\}$  is tight as  $t \rightarrow \infty$ , and collection of the sub sequential limits of  $\{\tilde{\xi}_t^n\}$ , denoted by  $\{\tilde{\mu}^n\}$ , is also tight. Let  $\tilde{\mu}$  be a sub-sequential limit of  $\{\tilde{\mu}^n\}$ . We claim that

$$\int_{\mathbb{R}^d} \left( \frac{1}{2} \sum_{i,j=1}^d a_{ij} \partial_{ii} f(x) + b(x, u_\varepsilon(x)) \cdot \nabla f(x) \right) \tilde{\mu}(dx) = 0, \quad \forall f \in \mathcal{C}_c^2(\mathbb{R}^d). \quad (65)$$

In fact, the claim follows from a similar argument as in Theorem 5.1. (65) shows that  $\tilde{\mu}$  is the unique invariant measure corresponds to the Markov control  $u_\varepsilon$ . Thus to complete the proof we only need to show that

$$\lim_{n \rightarrow \infty} J(\hat{X}^n(0), \hat{Z}^n) = \int_{\mathbb{R}^d} \tilde{r}(x, u_\varepsilon(x)) \tilde{\mu}(dx).$$

In view of Lemma 5.1, to show (64) it is enough to show that for any  $\psi \in \mathcal{C}_c(\mathbb{R}^d)$ , we have

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}^d} \psi(x) \tilde{r}(x, u^n(x)) \tilde{\mu}^n(dx) = \int_{\mathbb{R}^d} \psi(x) \tilde{r}(x, u_\varepsilon(x)) \tilde{\mu}(dx), \quad (66)$$

where

$$u_{ij}^n(x) x_j^- = \frac{1}{\sqrt{n}} \lfloor \sqrt{n} u_\varepsilon(x) x_j^- \rfloor, \quad \text{for } i \neq j.$$

Since  $\tilde{\mu}^n \rightarrow \tilde{\mu}$  and  $|\tilde{r}(x, u^n(x)) - \tilde{r}(x, u_\varepsilon(x))| \rightarrow 0$ , as  $n \rightarrow \infty$ , uniformly on compacts, (66) follows.  $\square$

We conclude the article with two important remarks.

**Remark 5.1.** It is evident from the arguments that one can replace  $r(\hat{Q}^n)$  by  $r(\hat{Q}^n) + h(\hat{Z}^n)$  for some convex function  $h: \mathbf{M} \rightarrow \mathbb{R}_+$  that lies in  $\mathcal{C}_{\text{pol}}(\mathbb{R}^d)$ . Since  $Z^n$  contains information about the idle times and the customers in service, one may want to minimize the cumulative idle time by associating such cost. For instance, if the service rate  $\mu_{ij}^n$  (the rate at which class- $i$  customers are served at station  $j$ ) is very small then it is logical to add a cost of type  $h(\hat{Z}^n)$  with a high payoff associated to  $Z_{ij}^n$ .

**Remark 5.2.** One might also put additional constraints on the service mechanism so that certain class of customers do not get served by some pools. The arguments of this article still go through in that case. Note that under such constraints we have  $i \nrightarrow j$  for some  $i, j$  with  $i \neq j$ . The only required change in such case is to restrict the corresponding entries of the matrices in  $\mathbf{M}$  to 0 whenever  $i \nrightarrow j$  for  $i \neq j$ .

## Acknowledgments

The author would like to thank the Editor and the anonymous referees for their helpful comments in improving the article. Author is grateful to Prof. Rami Atar for his comments on this project. Thanks are due to Prof. Kavita Ramanan for suggesting references related to this project.

## References

- Aliprantis CD, Border KC (2006) *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Third ed. (Springer, Berlin). [MR2378491](#).
- Arapostathis A, Pang G (2016) Ergodic diffusion control of multiclass multi-pool networks in the Halfin-Whitt regime. *Ann. Appl. Probab.* 26(5):3110–3153. [MR3563203](#)
- Arapostathis A, Pang G (2017) Infinite horizon average optimality of the  $N$ -network queueing model in the Halfin-Whitt regime. Preprint [arXiv:1602.03275](#).
- Arapostathis A, Biswas A, Pang G (2015) Ergodic control of multi-class  $M/M/N + M$  queues in the Halfin-Whitt regime. *Ann. Appl. Probab.* 25(6):3511–3570. [MR3404643](#).
- Arapostathis A, Borkar VS, Ghosh MK (2012) *Ergodic Control of Diffusion Processes*. Encyclopedia of Mathematics and Its Applications, Vol. 143 (Cambridge University Press, Cambridge, UK).
- Armony M (2005) Dynamic routing in large-scale service systems with heterogeneous servers. *Queueing Syst.* 51(3–4):287–329. [MR2189596](#).
- Atar R (2005) Scheduling control for queueing systems with many servers: Asymptotic optimality in heavy traffic. *Ann. Appl. Probab.* 15(4):2606–2650.
- Atar R, Biswas A (2014) Control of the multiclass  $G/G/1$  queue in the moderate deviation regime. *Ann. Appl. Probab.* 24(5):2033–2069. [MR3226171](#).
- Atar R, Giat C, Shimkin N (2010) The  $c\mu/\theta$  rule for many-server queues with abandonment. *Oper. Res.* 58(5):1427–1439.
- Atar R, Giat C, Shimkin N (2011) On the asymptotic optimality of the  $c\mu/\theta$  rule under ergodic cost. *Queueing Syst.* 67(2):127–144.
- Atar R, Mandelbaum A, Reiman MI (2004) Scheduling a multi class queue with many exponential servers: Asymptotic optimality in heavy traffic. *Ann. Appl. Probab.* 14(3):1084–1134.
- Atar R, Mandelbaum A, Shaikh G (2009) Simplified control problems for multiclass many-server queueing systems. *Math. Oper. Res.* 34(4):795–812. [MR2573496](#).
- Biswas A (2014) Risk-sensitive control for the multiclass many-server queues in the moderate deviation regime. *Math. Oper. Res.* 39(3):908–929. [MR3247010](#).
- Biswas A, Ishii H, Saha S, Wang L (2017) On viscosity solution of hjb equations with state constraints and reflection control. *SIAM J. Control Optim.* 55(1):365–396.
- Budhiraja A, Ghosh AP, Lee C (2011) Ergodic rate control problem for single class queueing networks. *SIAM J. Control Optim.* 49(4):1570–1606. [MR2817491](#).
- Dai JG, Tezcan T (2008) Optimal control of parallel server systems with many servers in heavy traffic. *Queueing Syst.* 59(2):95–134.
- Dai JG, Tezcan T (2011) State space collapse in many-server diffusion limits of parallel server systems. *Math. Oper. Res.* 36(2):271–320. [MR2828761](#).
- Gilbarg D, Trudinger NS (1983) *Elliptic Partial Differential Equations of Second Order*. Grundlehren der Mathematischen Wissenschaften, Second ed., Vol. 224 (Springer, Berlin).
- Gurvich I, Whitt W (2009) Queue-and-idleness-ratio controls in many-server service systems. *Math. Oper. Res.* 34(2):363–396. [MR2554064](#).
- Harrison JM, Zeevi A (2004) Dynamic scheduling of a multiclass queue in the Halfin-Whitt heavy traffic regime. *Oper. Res.* 52(2):243–257.
- Ichihara N, Sheu S-J (2013) Large time behavior of solutions of Hamilton-Jacobi-Bellman equations with quadratic nonlinearity in gradients. *SIAM J. Math. Anal.* 45(1):279–306.
- Iravani SMR, Kolfal B, Oyen MPV (2007) Call-center labor cross-training: It's a small world after all. *Management Sci.* 53(7):1102–1112.
- Kallenberg O (2002) *Foundations of Modern Probability*. Probability and Its Applications, Second ed. (Springer, New York).
- Kim J, Ward AR (2013) Dynamic scheduling of a  $GI/GI/1 + GI$  queue with multiple customer classes. *Queueing Syst.* 75(2–4):339–384.
- Krantz SG (2001) *Function Theory of Several Complex Variables* (AMS Chelsea Publishing, Providence, RI), RI Reprint of the 1992 edition. [MR1846625](#).
- Mandelbaum A, Stolyar AL (2004) Scheduling flexible servers with convex delay costs: Heavy-traffic optimality of the generalized  $c\mu$ -rule. *Oper. Res.* 52(6):836–855.
- Pang G, Talreja R, Whitt W (2007) Martingale proofs of many-server heavy-traffic limits for Markovian queues. *Probab. Surv.* 4:193–267.
- Ramanan K, Reiman MI (2003) Fluid and heavy traffic diffusion limits for a generalized processor sharing model. *Ann. Appl. Probab.* 13(1):100–139. [MR1951995](#).
- Ramanan K, Reiman MI (2008) The heavy traffic limit of an unbalanced generalized processor sharing model. *Ann. Appl. Probab.* 18(1):22–58. [MR2380890](#).
- Sznajder R, Filar JA (1992) Some comments on a theorem of Hardy and Littlewood. *J. Optim. Theory Appl.* 75(1):201–208. [MR1189274](#).
- van Mieghem JA (1995) Dynamic scheduling with convex delay costs: The generalized  $c\mu$  rule. *Ann. Appl. Probab.* 5(3):809–833.
- Ward AR, Armony M (2013) Blind fair routing in large-scale service systems with heterogeneous customers and servers. *Oper. Res.* 61(1):228–243. [MR3042753](#).
- Yosida K (1980) *Functional Analysis*. Grundlehren der Mathematischen Wissenschaften, Sixth ed., Vol. 123 (Springer, Berlin).