

THE LARGE DEVIATION PRINCIPLE FOR A GENERAL CLASS OF QUEUEING SYSTEMS I

PAUL DUPUIS AND RICHARD S. ELLIS

ABSTRACT. We prove the existence of a rate function and the validity of the large deviation principle for a general class of jump Markov processes that model queueing systems. A key step in the proof is a local large deviation principle for tubes centered at a class of piecewise linear, continuous paths mapping $[0, 1]$ into \mathbb{R}^d . In order to prove certain large deviation limits, we represent the large deviation probabilities as the minimal cost functions of associated stochastic optimal control problems and use a subadditivity-type argument. We give a characterization of the rate function that can be used either to evaluate it explicitly in the cases where this is possible or to compute it numerically in the cases where an explicit evaluation is not possible.

1. INTRODUCTION

While the large deviation literature on Markov processes is extensive, the vast majority of papers study processes having the property that the components of their generators depend smoothly upon the spatial parameter. The present paper is the latest in a series that analyze large deviation phenomena for Markov processes violating this smoothness condition. We refer to these processes as having “discontinuous statistics.” Our first paper [7] proves a general large deviation upper bound for a wide class of such processes. In our paper [5] and in Chapter 8 of our forthcoming book [6] the large deviation principle is proved for random walk models with discontinuous statistics that are natural generalizations of Cramér’s Theorem and Mogulskii’s Theorem [11], respectively. In the present paper we prove the existence of a rate function and the validity of the large deviation principle for a general class of jump Markov processes with discontinuous statistics that model queueing systems. We can handle a wide variety of problems with different geometries, including problems with finite buffers (see Example 4.4). Because of the generality of the model that we introduce, it should come as no surprise that an explicit formula for the rate function is in many cases not available. In a sequel to this paper, using the formalism

Received by the editors February 14, 1994.

1991 *Mathematics Subject Classification.* Primary 60F10, 60K25; Secondary 93E20.

Research of the first author was supported in part by the National Science Foundation (NSF-DMS-9115762), Air Force Office of Scientific Research (F49620-93-1-0264), and Army Research Office (DAAH04-93-G-0070).

Research of the second author was supported in part by the National Science Foundation (NSF-DMS-9123575).

developed here, we will derive a number of properties of the rate function and give an explicit formula for it in several cases of interest.

To date the large deviation principle has been proved for a variety of queueing systems having special structures. In the case of tandem queues, contraction mapping techniques and the Skorohod problem have been used. These techniques originated in [1], were extended to domains with corners in [4, 8], and have been applied to queueing systems in [12, 14]. Using partial differential equation methods and in particular viscosity solutions, the paper [9] proves the large deviation principle for a class of product form networks, which includes Jackson networks. In addition, the results of Chapter 8 of [6] allow one to treat a general class of stable queueing systems in two dimensions that are modeled by discrete-time Markov chains [10]. In terms of generality the present paper is a vast improvement. We prove the large deviation principle for a very wide class of lattice-based queueing systems in arbitrary dimensions under the assumption that the queueing systems are modeled by jump Markov processes. As in our forthcoming book [6], we make important use of ideas from stochastic optimal control theory, representing the large deviation probabilities as the minimal cost functions of associated stochastic optimal control problems. The main new feature with which we must deal is the geometry of the problem. The methods of this paper are adaptable in a straightforward way to a class of Markov-modulated and related non-Markovian processes. These extensions, as well as the harder extensions to more general state spaces, are currently under investigation.

We now turn to our earlier work concerning Markov processes with discontinuous statistics. Our first paper [7] proves a large deviation upper bound for a general class of locally infinitely divisible Markov processes taking values in \mathbb{R}^d , the components of the generators of which are allowed to have essentially arbitrary discontinuities. This class includes the jump Markov processes modeling queueing systems that are treated in the present paper. Our second paper [5] and Chapter 8 of [6] show that in the presence of special geometries the upper bound in [7] can be improved. In fact, these papers prove the full large deviation principle. We will now discuss the main result in Chapter 8 of [6] because it generalizes the setup of [5] and is closely related to the contents of the present work.

Consider the sets

$$\Lambda^{(1)} \doteq \{x \in \mathbb{R}^d : x_1 \leq 0\}, \quad \Lambda^{(2)} \doteq \{x \in \mathbb{R}^d : x_1 > 0\},$$

and

$$\partial \doteq \{x \in \mathbb{R}^d : x_1 = 0\},$$

where x_1 denotes the first component of $x \in \mathbb{R}^d$. Let $\{\mu_x^{(1)}, x \in \mathbb{R}^d\}$ and $\{\mu_x^{(2)}, x \in \mathbb{R}^d\}$ be two families of probability measures on \mathbb{R}^d such that for each $i = 1, 2$ the function mapping $x \in \mathbb{R}^d \mapsto \mu_x^{(i)}$ is continuous in the topology of weak convergence. For $n \in \mathbb{N}$, $j \in \{0, 1, \dots, n-1\}$, and $x \in \mathbb{R}^d$ we define random vectors $X_0^n = x$ and

$$X_{j+1}^n = X_j^n + \frac{1}{n} Y_j^n$$

and assume that Y_j^n has the following conditional distribution:

$$P_n\{Y_j^n \in dy | X_0^n, X_1^n, \dots, X_j^n\} \doteq \begin{cases} \mu_\xi^{(1)}(dy) & \text{if } X_j^n = \xi \in \Lambda^{(1)}, \\ \mu_\xi^{(2)}(dy) & \text{if } X_j^n = \xi \in \Lambda^{(2)}. \end{cases}$$

Let $\{\xi^n, n \in \mathbb{N}\}$ be a sequence of points in the open halfspace $\Lambda^{(2)}$ converging to a point $\xi \in \partial$, which is a subset of the closed halfspace $\Lambda^{(1)}$. The discontinuous statistics arise since we do not assume that the sequence of measures $\{\mu_{\xi^n}^{(2)}, n \in \mathbb{N}\}$ converges weakly to the measure $\mu_\xi^{(1)}$. We now introduce the piecewise linear interpolation of the Markov chain $\{X_j^n, j = 0, 1, \dots, n\}$, defined by the formula

(1.1)

$$X^n(t) \doteq X_j^n + \left(t - \frac{j}{n}\right) Y_j^n \text{ for } t \in [j/n, (j+1)/n], j = 0, 1, \dots, n-1.$$

These processes take values in the space $\mathcal{E}([0, 1] : \mathbb{R}^d)$. The main theorem in Chapter 8 of [6] is the large deviation principle for the random walks $\{X^n(t), t \in [0, 1], n \in \mathbb{N}\}$.

In the context of the present paper it is worthwhile to examine the formula for the rate function given in Chapter 8 of [6]. Although it is given explicitly, its complicated form illuminates the difficulty in obtaining an analogous explicit formula for the more complicated geometries that arise in the study of queueing systems. For $i = 1, 2, \xi \in \mathbb{R}^d$, and $\beta \in \mathbb{R}^d$, we let $L^{(i)}(\xi, \beta)$ be the Cramér function associated with the family of probability measures $\{\mu_x^{(i)}, x \in \mathbb{R}^d\}$:

$$L^{(i)}(\xi, \beta) \doteq \sup_{\alpha \in \mathbb{R}^d} \left\{ \langle \alpha, \beta \rangle - \log \int_{\mathbb{R}^d} \exp\langle \alpha, y \rangle \mu_\xi^{(i)}(dy) \right\}.$$

The symbol $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product on \mathbb{R}^d . For $\xi \in \mathbb{R}^d$ and $\beta \in \mathbb{R}^d$ we define the function

$$L^{(0)}(\xi, \beta) \doteq \inf \left\{ \rho^{(1)} L^{(1)}(\xi, \beta^{(1)}) + \rho^{(2)} L^{(2)}(\xi, \beta^{(2)}) \right\},$$

where the infimum is taken over quantities $\rho^{(1)} \in \mathbb{R}, \rho^{(2)} \in \mathbb{R}, \beta^{(1)} \in \mathbb{R}^d$, and $\beta^{(2)} \in \mathbb{R}^d$ satisfying

(1.2) $\rho^{(1)} + \rho^{(2)} = 1, \rho^{(1)} \geq 0, \rho^{(2)} \geq 0,$

(1.3) $(\beta^{(1)})_1 \geq 0, (\beta^{(2)})_1 \leq 0,$

(1.4) $\rho^{(1)} \beta^{(1)} + \rho^{(2)} \beta^{(2)} = \beta.$

The subscripts 1 denote the one-components of the vectors. We then define for $\xi \in \mathbb{R}^d$ and $\beta \in \mathbb{R}^d$

(1.5)
$$\tilde{L}(\xi, \beta) = \begin{cases} L^{(1)}(\xi, \beta) & \text{if } \xi_1 < 0, \\ L^{(0)}(\xi, \beta) & \text{if } \xi_1 = 0, \\ L^{(2)}(\xi, \beta) & \text{if } \xi_1 > 0. \end{cases}$$

Under appropriate conditions given in Chapter 8 of [6] we obtain the rate function for the random walk model. For $x \in \mathbb{R}^d$, which denotes the initial point $X^n(0)$ for each $n \in \mathbb{N}$, the rate function is given by the formula

$$(1.6) \quad I_x(\varphi) \doteq \int_0^1 \tilde{L}(\varphi(t), \dot{\varphi}(t)) dt$$

whenever $\varphi \in \mathcal{C}([0, 1] : \mathbb{R}^d)$ is absolutely continuous and satisfies $\varphi(0) = x$. For all other $\varphi \in \mathcal{C}([0, 1] : \mathbb{R}^d)$, we set $I_x(\varphi) \doteq \infty$.

The boundary ∂ influences the rate function I_x through the function $L^{(0)}(\xi, \beta)$. Through equation (1.3) the form of $L^{(0)}(\xi, \beta)$ incorporates a necessary and sufficient condition for the stability of the one-component of a class of controlled random walks. The necessary condition is for the large deviation upper bound and the sufficient condition for the large deviation lower bound. For example, with reference to the latter bound let us consider the problem of estimating the probability that $\sup_{t \in [0, 1]} \|X^n(t) - \beta t\| < \varepsilon$, where $\varepsilon > 0$ and β is a vector in \mathbb{R}^d satisfying $\beta_1 = 0$. In terms of the classical change of measure argument one would consider changes of measures that shift the drifts of the processes in the halfspaces $\Lambda^{(1)}$ and $\Lambda^{(2)}$ to $\beta^{(1)}$ and $\beta^{(2)}$, respectively, in such a way that the event of interest has a probability of order 1 as $n \rightarrow \infty$. However, unless the constraints in formula (1.3) are satisfied, the probability of the event is of order 0 as $n \rightarrow \infty$. There is an analogous interpretation of the constraints in (1.3) under the stochastic optimal control formulation of the problem in Chapter 8 of [6]. As explained in [5, Lemma 5.2], the stability constraints in (1.3) are needed in a related fashion in order to prove the large deviation upper bound. We note that formula (1.3) is *not* a stability condition on the original process but rather on the process obtained from the original process by a change of measure.

The processes that will be studied in the present paper as models of queueing systems are lattice-based jump Markov processes with discontinuous statistics that generalize the type of discontinuity considered in [5] and in Chapter 8 of [6]. The present processes are defined in terms of much more complicated geometries involving three or more regions of constant (or smooth) statistical behavior separated by an arbitrary finite number of intersecting hyperplanes of codimension one across which the statistical behavior can change discontinuously. In the special case of two halfspaces one can show that the jump Markov processes are superexponentially close to a sequence of processes of the form (1.1) and that because of this the rate function for the corresponding queueing system takes the same form as in equations (1.3)–(1.6). In the case of the general queueing processes considered here, the difficulty in determining an explicit formula for the rate function along intersections of two or more hyperplanes arises from the extremely hard problem of determining necessary and sufficient conditions for stability that are analogous to formula (1.3) and are valid for this more complicated geometry. However, when additional structure is present—for example, if the original network is a Jackson network—this problem can in some cases be resolved.

In order to motivate the notation in this paper and to suggest the variety of situations that we can handle, we give in Section 2 a detailed description of a

special but interesting two-dimensional queueing system that we call the “modified join-the-shorter-queue” model. In subsequent sections we generalize this example to a wide class of queueing systems having as state space a subset \mathcal{S} of the d -dimensional integer lattice \mathbb{Z}^d , where $d \geq 2$. They may incorporate a variety of input disciplines and service disciplines together with other features such as feedback.

In the general setup described in Section 3, the assumptions are reasonably weak. Condition 3.4 is a communication condition, which is satisfied for example by any tandem queue as well as by the modified join-the-shorter queue model of Section 2. Condition 3.1 has three components: that the state space \mathcal{S} is partitioned into finitely many subsets called facets, that the jump intensities of the jump Markov processes modeling the queueing systems are constant in each facet, and that only finitely many of the jump intensities are positive. In Section 4 we consider queueing models for which the jump intensities have some additional state dependence. Namely, we allow discontinuous statistics that are defined by a state dependent jump intensity function $r(x, v)$ that is continuous, and not necessarily constant, within each facet. The processes of Section 4 include the processes of Section 3 as a special case.

Theorem 4.3 states the large deviation principle for the jump Markov processes of Section 4. Although we defer to that section the detailed characterization of the rate function, it is worthwhile to outline here the main steps in the proof.

- Proposition 3.7 obtains local large deviation estimates for “tubes” centered at paths mapping $[0, 1]$ into \mathbb{R}^d and having constant velocities. These paths have the form $\{x + \beta t, t \in [0, 1]\}$.
- Proposition 5.1 uses Proposition 3.7 and the Markov property to deduce a local large deviation principle for tubes centered at a class of piecewise linear, continuous paths mapping $[0, 1]$ into \mathbb{R}^d . For such paths φ with starting points x we obtain a local rate function $I_x(\varphi)$.
- The rate function in the large deviation principle stated in Theorem 4.3 is the lower semicontinuous regularization of the function I_x obtained in the previous step.

Proposition 3.7 is proved for systems having jump intensities that are constant within each facet of the state space. In the proof suitable large deviation probabilities are represented in terms of the minimal cost functions of associated stochastic optimal control problems, and a subadditivity-type argument is used to show that the minimal cost functions satisfy appropriate limits. This is one of the main contributions of the paper. Proposition 5.1 is proved by “localizing” the jump intensities, obtaining jump intensities that are constant on each facet and thus to which Proposition 3.7 can be applied. The localization argument is given in Section 4 because it is needed to define the integrand of the local rate function I_x . The large deviation principle stated in Theorem 4.3 is proved in Section 5, and the local large deviation principle stated in Proposition 5.1 is proved in Section 6.

The present paper shares some features with [6]. Like there, a key step in the proof of the large deviation principle is to represent the large deviation probabilities as the minimal cost functions of a sequence of stochastic optimal control

problems. However, rather than use weak convergence methods to calculate the limits of the minimal cost functions as we do in [6], here the existence of the limits is shown using a subadditivity-type argument. Our experience in [6] leads us to think that the weak convergence methods are convenient for treating problems in which an explicit formula for the rate function is available. Because in the generality presented in this paper an explicit formula for the rate function is in many cases not available, the weak convergence methods of [6] are not used. We note that the stochastic control representation and weak convergence methods can be used to identify the rate function when identification is possible and can be used to calculate the rate function numerically when identification is not possible. These topics will be taken up in the sequel to this paper. Although we will not enumerate such extensions, the general approach to the large deviation analysis of queueing systems adopted in this paper can be extended in many directions to cover numerous related queueing models.

2. EXAMPLE: MODIFIED JOIN-THE-SHORTER-QUEUE MODEL

In order to motivate the notation in this paper and to suggest the variety of queueing systems that we can handle, we describe a two-dimensional queueing system that is a special case of the general model considered in Section 3. We refer to it as the “modified join-the-shorter-queue” model.

We consider a queueing system with two queues, three independent input streams, and one class of customers. The first input stream feeds into queue 1 at the rate $\alpha > 0$, the second input stream feeds into queue 2 at the rate $\beta > 0$, and the third input stream joins the shorter of the two queues at the rate $\gamma > 0$. If both queues are equal at the moment of arrival of a customer in the third stream, then by tossing an independent fair coin the customer decides which of the two queues to enter. In addition queues 1 and 2 serve customers at the rates $\delta > 0$ and $\mu > 0$, respectively. The system is represented as in Figure 1.

We model the system by a jump Markov process $\{X(t), t \in [0, \infty)\}$, where $X(t) = (X_1(t), X_2(t))$ and for each $i \in \{1, 2\}$ $X_i(t)$ denotes the sum of the number of customers waiting in queue i and the number of customers being served in queue i . The state space of $X(t)$ is the nonnegative quadrant

$$\mathcal{S} \doteq \{(j, k) \in \mathbb{Z}^2 : j \geq 0, k \geq 0\}$$

of the two-dimensional integer lattice \mathbb{Z}^2 . At each point in the interior of \mathcal{S} the possible jump directions are the four nearest neighbors. At each point in the boundary of \mathcal{S} the possible jump directions are those nearest neighbors lying in \mathcal{S} . The jump intensities at each point in \mathcal{S} depend on where in \mathcal{S} the point lies.

In order to give a complete description, we introduce six subsets of \mathbb{R}^2 that will lead to a partition of the state space \mathcal{S} . In preparation for the introduction of the general queueing systems in Section 3, let us label the subsets using a notation that is suited for this general case. Consider the vectors $w^{(1)} \doteq (1, 0)$, $w^{(2)} \doteq (0, 1)$, and $w^{(3)} \doteq (1, -1)$. For each $i \in \{1, 2, 3\}$ let \circ_i represent one of the symbols $>, =, <$. We label each facet by the notation $\mathcal{F}^{(\circ_1, \circ_2, \circ_3)}$ to

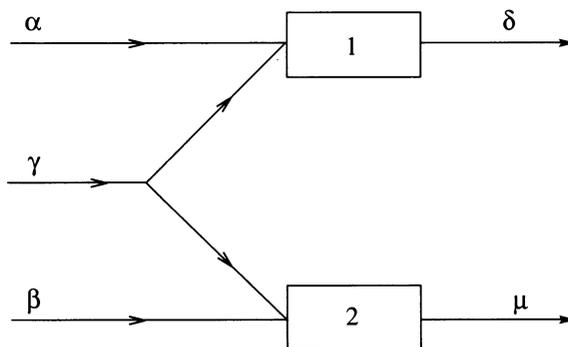


FIGURE 1. The modified join-the-shorter-queue model.

indicate that the facet is the intersection of the three sets $\{y \in \mathbb{R}^2 : \langle y, w^{(i)} \rangle \circ_i 0\}$. In other words,

$$\mathcal{F}^{(\circ_1, \circ_2, \circ_3)} \doteq \{y \in \mathbb{R}^2 : \langle y, w^{(i)} \rangle \circ_i 0 \text{ for all } i = 1, 2, 3\}.$$

In terms of this notation we have

$$\mathcal{F}^{(>, >, >)} \doteq \{y \in \mathbb{R}^2 : y_1 > y_2 > 0\},$$

$$\mathcal{F}^{(>, >, <)} \doteq \{y \in \mathbb{R}^2 : y_2 > y_1 > 0\},$$

$$\mathcal{F}^{(>, >, =)} \doteq \{y \in \mathbb{R}^2 : y_1 = y_2 > 0\},$$

$$\mathcal{F}^{(>, =, >)} \doteq \{y \in \mathbb{R}^2 : y_1 > 0, y_2 = 0\},$$

$$\mathcal{F}^{(=, >, <)} \doteq \{y \in \mathbb{R}^2 : y_1 = 0, y_2 > 0\},$$

$$\mathcal{F}^{(=, =, =)} \doteq \{y \in \mathbb{R}^2 : y_1 = y_2 = 0\}.$$

Each of the other sets of the form $\mathcal{F}^{(\circ_1, \circ_2, \circ_3)}$ has an empty intersection with the state space \mathcal{S} .

For each of the sets in the last display, we refer to the nonempty intersection $\mathcal{F}^{(\circ_1, \circ_2, \circ_3)} \cap \mathcal{S}$ as a *facet* of \mathcal{S} and use the same term when referring to the sets $\mathcal{F}^{(\circ_1, \circ_2, \circ_3)}$ themselves. The first and second facets are each a “wedge” of dimension 2; the third, fourth, and fifth facets are each a halfline of dimension 1; and the sixth facet is a vertex of dimension 0. Within each facet the jump intensities are constant but they differ from facet to facet. The jump intensities for any point in \mathcal{S} are represented in Figure 2.

This modified join-the-shorter-queue model satisfies both of the conditions to be introduced in Section 3. In particular, the communication condition expressed in Condition 3.4 is satisfied because any point $(a, b) \in \mathcal{S}$ can be reached with positive probability from any point $(j, k) \in \mathcal{S}$ by a series of

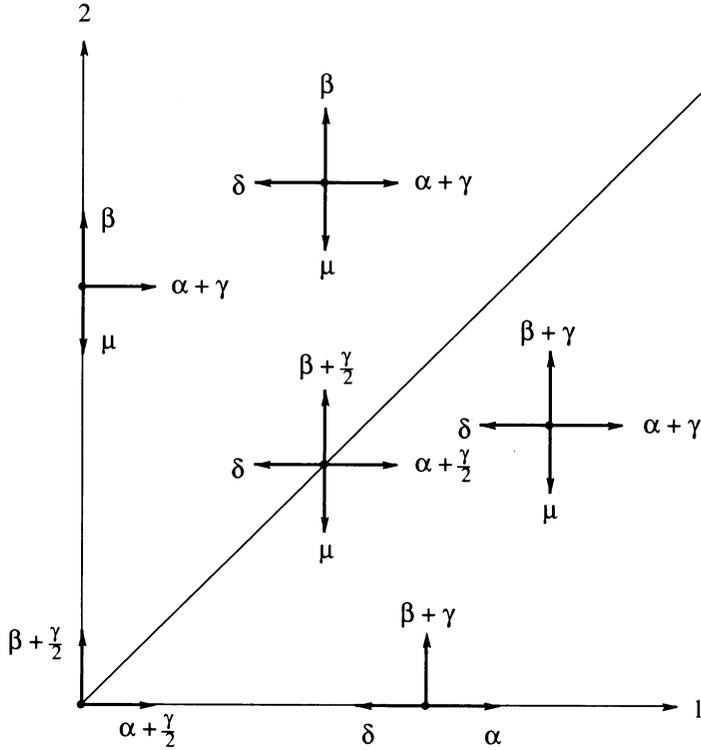


FIGURE 2. Jump intensities for the modified join-the-shorter-queue model.

$(|a - j| + |b - k|)$ jumps. This is not the case for the standard join-the-shorter-queue model. In the standard model, for $(j, k) \in (\mathcal{F}(>, >, >) \cup \mathcal{F}(>, =, >)) \cap \mathcal{S}$ the jump to the point $(j + 1, k)$ is suppressed, and for $(j, k) \in (\mathcal{F}(>, >, <) \cup \mathcal{F}(=, >, <)) \cap \mathcal{S}$ the jump to the point $(j, k + 1)$ is suppressed. In order to analyze the standard join-the-shorter-queue model, one must work with a suitable weakening of the communication condition. The relationship between the modified join-the-shorter-queue model and the standard join-the-shorter-queue model is analogous to the relationship between a nondegenerate diffusion process and a degenerate diffusion process. Unfortunately, such degenerate queueing systems seem at the present time to require a case-by-case treatment.

3. DEFINITION OF THE LOCAL RATE FUNCTION

This section consists of two subsections. In the first we introduce a class of queueing systems. They are defined in terms of a state space, a partition of the state space into sets called facets, jump intensities that are constant in each

facet, and a jump Markov process having these jump intensities. In the second subsection we begin our study of these queueing systems, introducing a rate function for a certain local large deviation principle. In order to prove this, we represent the large deviation probabilities as the minimal cost functions of associated stochastic optimal control problems (Theorem 3.3). The large deviation limits are proved using this representation formula and a subadditivity-type argument. In Section 4 we will introduce a wider class of queueing systems in which the jump intensities can have some additional state dependence. In order to apply the material of Section 3 to these systems, we will have to carry out a local analysis in the neighborhood of each point in the state space. This local analysis will replace the original jump Markov process by a process having the form given in Subsection 3.1.

3.1. A description of the local model. We consider a queueing system modeled by a stationary jump Markov process. In order to describe the process, we take as given the following data:

- The state space \mathcal{S} of the process, which is a subset of \mathbb{Z}^d and contains the origin 0.
- The jump intensities at each point $x \in \mathcal{S}$. These are of the form $\{r(x, v), v \in \mathbb{Z}^d\}$, where $r(x, v) \geq 0$ for all x and v and where $r(x, v) > 0$ only if $x + v \in \mathcal{S}$. The quantity $r(x, v)$ gives the jump intensity from x to $x + v$.

We call $r(\cdot, \cdot)$ the jump intensity function. Additional assumptions on this function are given in Conditions 3.1 and 3.4.

Let $\{X(t), t \in [0, \infty)\}$ be a stationary jump Markov process with state space \mathcal{S} and jump intensities $\{r(x, v), x \in \mathcal{S}, v \in \mathbb{Z}^d\}$. For $x \in \mathcal{S}$ and f any bounded real-valued function on \mathbb{Z}^d these processes can be characterized by their generators, which are of the form

$$\mathcal{L}f(x) \doteq \sum_{v \in \mathbb{Z}^d} r(x, v) [f(x + v) - f(x)].$$

For $0 \leq t_1 < t_2 < \infty$ and A a subset of \mathbb{R}^d we define $\mathcal{D}([t_1, t_2] : A)$ to be the space of functions that map $[t_1, t_2]$ into A and that are continuous from the right and have limits from the left. The space $\mathcal{D}([t_1, \infty) : A)$ is defined similarly. The jump Markov process $\{X(t), t \in [0, \infty)\}$ takes values in $\mathcal{D}([0, \infty) : \mathcal{S})$.

The structure of the queueing systems to be considered in this section depends on a finite collection of nonzero vectors $\{w^{(i)}, i = 1, 2, \dots, \kappa\}$, where each $w^{(i)} \in \mathbb{R}^d$ and each $w^{(i)}$ has rational coordinates. This condition concerning rational coordinates is needed in order to establish formula (3.5). In terms of these vectors we define the facets of \mathbb{R}^d associated with the given model. The facets are defined by generalizing the notation introduced at the end of Section 2, where the modified join-the-shorter-queue was discussed. As we explain in Condition 3.1, the jump intensities are to be constant in the intersection of each facet with the state space \mathcal{S} .

Before defining the facets, let us anticipate an important issue concerning the main result in this paper, which is the large deviation principle stated in Theorem 4.3. Its proof relies on a localization argument based on a family of

jump Markov processes that are related to the original jump Markov process $\{X(t), t \in [0, \infty)\}$ but are defined in terms of vectors $w^{(i)}$ that in general are a proper subset of those just introduced. The setup in this and the next subsection is sufficiently broad to cover all these processes.

We return to the main development. For each $i \in \{1, 2, \dots, \kappa\}$ let \circ_i represent one of the three symbols $>, =, <$. We label each facet by the notation $\mathcal{F}^{(\circ_1, \dots, \circ_\kappa)}$ to indicate that the facet is the intersection of the sets $\{y \in \mathbb{R}^d : \langle y, w^{(i)} \rangle \circ_i 0\}$. In other words,

$$(3.1) \quad \mathcal{F}^{(\circ_1, \dots, \circ_\kappa)} \doteq \{y \in \mathbb{R}^d : \langle y, w^{(i)} \rangle \circ_i 0 \text{ for all } i = 1, 2, \dots, \kappa\}.$$

Note that if $y \in \mathcal{F}^{(\circ_1, \dots, \circ_\kappa)}$, then $\alpha y \in \mathcal{F}^{(\circ_1, \dots, \circ_\kappa)}$ for every $\alpha > 0$. Thus each facet $\mathcal{F}^{(\circ_1, \dots, \circ_\kappa)}$ is a cone. Assumptions concerning these sets and the jump intensity function $r(x, v)$ are given in the next condition. Because of the structure of the state space \mathcal{S} as expressed in part (a) of Condition 3.1, the assumption that the jump intensities are constant in each of the facets is a radial homogeneity condition. For each x the support of $r(x, \cdot)$ is defined to be the set

$$\Sigma(x) \doteq \{v \in \mathbb{Z}^d : r(x, v) > 0\}.$$

Condition 3.1. *We consider a finite collection of nonzero vectors $\{w^{(i)}, i = 1, 2, \dots, \kappa\}$ in \mathbb{R}^d such that each $w^{(i)}$ has rational components. We define the sets $\mathcal{F}^{(\circ_1, \dots, \circ_\kappa)}$ in terms of these vectors by equation (3.1). The sets $\{\mathcal{F}^{(\circ_1, \dots, \circ_\kappa)} \cap \mathcal{S}\}$ are a partition of \mathcal{S} . We refer to these sets as the facets of \mathcal{S} . For each sequence of symbols $(\circ_1, \dots, \circ_\kappa)$ we assume the following:*

- (a) *If $\mathcal{F}^{(\circ_1, \dots, \circ_\kappa)} \cap \mathcal{S} \neq \emptyset$, then $\mathcal{F}^{(\circ_1, \dots, \circ_\kappa)} \cap \mathbb{Z}^d \subset \mathcal{S}$.*
- (b) *$r(x, v)$ is independent of $x \in \mathcal{F}^{(\circ_1, \dots, \circ_\kappa)} \cap \mathcal{S}$.*
- (c) *For $x \in \mathcal{F}^{(\circ_1, \dots, \circ_\kappa)} \cap \mathcal{S}$ the support $\Sigma(x)$ consists of a finite number of vectors.*

As we discussed in the Introduction and in Section 2, examples of facets are “wedges,” halflines, and the faces of the nonnegative orthant in \mathbb{R}^d . Part (a) of Condition 3.1 states that if any point in such a set is contained in \mathcal{S} , then the entire intersection of that set with \mathbb{Z}^d is also contained in \mathcal{S} . In particular, all positive integral multiples of any point in $\mathcal{F}^{(\circ_1, \dots, \circ_\kappa)} \cap \mathcal{S}$ are also in \mathcal{S} .

There are a number of direct consequences of Condition 3.1 that we will need. For example, it implies that

$$(3.2) \quad r(\alpha x, v) = r(x, v) \text{ if } \alpha > 0 \text{ and } \alpha x \in \mathcal{S}.$$

It also implies that the cardinality of the support $\Sigma(x)$ has a bound that is independent of $x \in \mathcal{S}$. Hence there exists a positive integer B so that

$$(3.3) \quad \sup\{\|v\| : x \in \mathcal{S}, v \in \Sigma(x)\} \leq B.$$

In addition Condition 3.1 guarantees the existence of a constant $C \in [1, \infty)$ such that

$$(3.4) \quad 1/C \leq r(x, v) \leq C$$

for all $x \in \mathcal{S}$ and $v \in \mathbb{Z}^d$ satisfying $r(x, v) > 0$. For any $\alpha > 0$ we extend the definition of $r(\cdot, \cdot)$ from $\mathcal{S} \times \mathbb{Z}^d$ to $(\alpha\mathcal{S}) \times \mathbb{Z}^d$ by setting $r(\alpha x, v) \doteq r(x, v)$

whenever $x \in \mathcal{S}$ and $v \in \mathbb{Z}^d$. Because of the radial homogeneity of the jump intensity function, this extension is well defined. The extension will allow us to define $r(\cdot, \cdot)$ for certain scaled lattices that we introduce below.

For $n \in \mathbb{N}$ we consider the scaled state spaces

$$\mathcal{S}^n \doteq \left\{ \frac{1}{n} y : y \in \mathcal{S} \right\}$$

and let Γ be the linear subspace

$$\Gamma \doteq \{ \beta \in \mathbb{R}^d : \langle \beta, w^{(i)} \rangle = 0 \text{ for all } i = 1, 2, \dots, \kappa \} = \mathcal{F}^{(=, \dots, =)}.$$

Condition 3.1 and the fact that $0 \in \mathcal{S}$ imply that the scaled state spaces $\{ \mathcal{S}^n, n \in \mathbb{N} \}$ are ‘‘asymptotically dense’’ in Γ in the sense that

$$(3.5) \quad \Gamma = \{ \beta \in \mathbb{R}^d : \beta = \lim_{n \rightarrow \infty} \beta^n \text{ for some sequence } \{ \beta^n, n \in \mathbb{N} \} \\ \text{with each } \beta^n \in \Gamma \cap \mathcal{S}^n \}.$$

In Proposition 3.7 in the next subsection we will prove a local large deviation limit for vectors β lying in this set Γ .

3.2. A rate function for the local model. In terms of the local model of the previous subsection, we now introduce a sequence of scaled jump Markov processes for which the large deviation principle will be proved. We then represent the corresponding large deviation probabilities in terms of the minimal cost functions of a sequence of associated stochastic optimal control problems. In order to obtain the large deviation estimates, we will need a communication condition for the jump Markov processes that will be used as a controllability condition for the associated controlled jump Markov processes. This communication/controllability condition is given in Condition 3.4.

Let $\{X(t), t \in [0, \infty)\}$ be the jump Markov process with state space $\mathcal{S} \subset \mathbb{Z}^d$, starting point $X^n(0) = y \in \mathcal{S}^n$, and jump intensities $r(x, v)$ satisfying Condition 3.1. For $n \in \mathbb{N}$ we define the processes $\{X^n(t), t \in [0, 1]\}$ by the formula

$$X^n(t) = \frac{1}{n} X(nt).$$

For each fixed n this process is a stationary jump Markov process having state space \mathcal{S}^n and taking values in $\mathcal{D}([0, 1] : \mathcal{S}^n)$. Using equation (3.2), one easily verifies that for $y \in \mathcal{S}^n$ the generator of the process is given by the formula

$$\mathcal{L}^n f(y) = n \sum_{v \in \mathbb{Z}^d} r(y, v) [f(y + v/n) - f(y)],$$

where f is any bounded real-valued function on \mathcal{S}^n .

Let β be any vector in \mathbb{R}^d . For $n \in \mathbb{N}$, $y \in \mathcal{S}^n$, $t \in [0, 1]$, and $\varepsilon \in (0, 1)$, we define the conditional probability

$$p^n(y, t; \beta, \varepsilon) \doteq P_{y,t}^n \left\{ \sup_{s \in [t, 1]} \|X^n(s) - s\beta\| < \varepsilon \right\},$$

where $P_{y,t}^n$ denotes probability conditioned on $X^n(t) = y$. For our needs in the next section we only need to consider those vectors β lying in the subspace

Γ . Our immediate goal will be to prove the existence of a convex function L mapping Γ into $[0, \infty)$ such that

$$\begin{aligned}
 & \lim_{\varepsilon \rightarrow 0} \lim_{\delta \rightarrow 0} \liminf_{n \rightarrow \infty} \inf_{\{y \in \mathcal{S}^n : \|y\| \leq \delta\}} \left(-\frac{1}{n} \log p^n(y, 0; \beta, \varepsilon) \right) \\
 (3.6) \quad & = \lim_{\varepsilon \rightarrow 0} \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\{y \in \mathcal{S}^n : \|y\| \leq \delta\}} \left(-\frac{1}{n} \log p^n(y, 0; \beta, \varepsilon) \right) \\
 & = L(\beta).
 \end{aligned}$$

These limits will be proved in Proposition 3.7. They will lead to the statement of the large deviation principle in Section 4 and its proof in Sections 5 and 6. The quantity $L(\beta)$ is uniquely determined by these limits.

In order to prove the limits in formula (3.6), we introduce the functions

$$(3.7) \quad q^n(y, t; \beta, \varepsilon) \doteq -\frac{1}{n} \log p^n(y, t; \beta, \varepsilon),$$

where we set $-\log 0 \doteq \infty$. Theorem 3.3 will represent $q^n(y, t; \beta, \varepsilon)$ as the minimal cost function of a control problem. This representation of the large deviation probabilities $q^n(y, t; \beta, \varepsilon)$ is a key step in the proof of the limits (3.6).

We need several facts about the functions $p^n(y, t; \beta, \varepsilon)$. For f a bounded function mapping \mathcal{S}^n into \mathbb{R} , the formal adjoint $(\mathcal{L}^n)^*$ of \mathcal{L}^n is defined by the formula

$$(\mathcal{L}^n)^* f(y) \doteq -n \sum_{v \in \mathbb{Z}^d} r(y, v) [f(y + v/n) - f(y)].$$

Fix $n \in \mathbb{N}$, $\beta \in \mathbb{R}^d$, and $\varepsilon > 0$. The functions $p^n(y, t; \beta, \varepsilon)$ satisfy the Kolmogorov backward equations

$$\begin{aligned}
 (3.8) \quad & \frac{\partial p^n(y, t; \beta, \varepsilon)}{\partial t} = (\mathcal{L}^n)^* p^n(y, t; \beta, \varepsilon) \\
 & \text{for } (y, t) \in \mathcal{S}^n \times (0, 1) \text{ satisfying } \|y - t\beta\| < \varepsilon;
 \end{aligned}$$

the boundary condition

$$(3.9) \quad p^n(y, t; \beta, \varepsilon) = 0 \text{ for } (y, t) \in \mathcal{S}^n \times [0, 1] \text{ satisfying } \|y - t\beta\| \geq \varepsilon;$$

and the terminal condition

$$(3.10) \quad p^n(y, 1, \beta, \varepsilon) = 1 \text{ for } y \in \mathcal{S}^n \text{ satisfying } \|y - \beta\| < \varepsilon.$$

For fixed $y \in \mathcal{S}^n$ let $(t_1(y), t_2(y))$ denote the open interval of values $t \in \mathbb{R}$ that satisfy $\|y - t\beta\| < \varepsilon$. If $t_2(y) \leq 1$, then the boundary condition (3.9) is valid in the “forward” t -direction; i.e., it is enforced in the sense that

$$\lim_{t \uparrow t_2(y)} p^n(y, t; \beta, \varepsilon) = 0.$$

The function $p^n(y, t; \beta, \varepsilon)$ is allowed to be discontinuous at $t = t_1(y)$ if $t_1(y) \geq 0$. Since for each $y \in \mathcal{S}^n$ there is a fixed upper bound on the number of $v \in \mathbb{Z}^d$ for which $r(y, v) > 0$, the system (3.8) is just a finite system of

linear ordinary differential equations. It follows that the equations (3.8), (3.9), and (3.10) have a unique solution.

We adopt the convention that $[0, a] \doteq \emptyset$ if $a < 0$ and that $[b, 1] \doteq \emptyset$ if $b > 1$. Let \mathcal{B}^n be the set of all $y \in \mathcal{S}^n$ such that $p^n(y, t; \beta, \varepsilon) > 0$ for some $t \in [0, 1]$ and consider any $y \in \mathcal{B}^n$. It is easy to check that if $p^n(y, t; \beta, \varepsilon) > 0$, then $p^n(y, s; \beta, \varepsilon) > 0$ for all $s \in (t_1(y) \vee 0, t]$. Therefore, given $y \in \mathcal{B}^n$ there are values $\tilde{t}_1(y)$ and $\tilde{t}_2(y)$ satisfying $t_1(y) \leq \tilde{t}_1(y) < \tilde{t}_2(y) \leq t_2(y)$ such that

$$p^n(y, t; \beta, \varepsilon) > 0 \text{ if } t \in (\tilde{t}_1(y), \tilde{t}_2(y)) \cap [0, 1]$$

and

$$p^n(y, t; \beta, \varepsilon) = 0 \text{ if } t \in [0, \tilde{t}_1(y)] \text{ or } t \in [\tilde{t}_2(y), 1].$$

In general we can take $\tilde{t}_1(y) = t_1(y)$, although in general we may have $\tilde{t}_2(y) < t_2(y)$.

We now write down the equations satisfied by the functions

$$q^n(y, t; \beta, \varepsilon) \doteq -\frac{1}{n} \log p^n(y, t; \beta, \varepsilon).$$

For $y \in \mathcal{B}^n$ and $t \in (\tilde{t}_1(y), \tilde{t}_2(y)) \cap (0, 1)$, we have the system of ordinary differential equations

$$(3.11) \quad \frac{\partial q^n(y, t; \beta, \varepsilon)}{\partial t} = \sum_{v \in \mathbb{Z}^d} r(y, v) (\exp(-n[q^n(y + v/n, t; \beta, \varepsilon) - q^n(y, t; \beta, \varepsilon)]) - 1),$$

where we use the convention that $\exp(-\infty) \doteq 0$. We also have

$$(3.12) \quad q^n(y, t; \beta, \varepsilon) = \infty \text{ for } (y, t) \in (\mathcal{S}^n \setminus \mathcal{B}^n \times [0, 1]) \cup (\mathcal{B}^n \times ([0, \tilde{t}_1(y)] \cup [\tilde{t}_2(y), 1])),$$

and the terminal condition

$$(3.13) \quad q^n(y, 1; \beta, \varepsilon) = 0 \text{ for } y \in \mathcal{S}^n \text{ satisfying } \|y - \beta\| < \varepsilon.$$

3.3. A stochastic control representation. In Theorem 3.3 we introduce a stochastic control representation for the functions $q^n(y, t; \beta, \varepsilon)$. This representation will be used to establish the existence of the limits in formula (3.6). The stochastic control representation consists of a controlled process and a cost structure.

Controlled process. Fix $n \in \mathbb{N}$. A measurable function $u^n(y, v, t)$ mapping $\mathcal{S}^n \times \mathbb{Z}^d \times [0, 1]$ into $[0, \infty)$ is said to be a control. It will be sufficient to consider only controls with the property that for $(y, v, t) \in \mathcal{S}^n \times \mathbb{Z}^d \times [0, 1]$

$$(3.14) \quad u^n(y, v, t) > 0 \text{ implies } r(y, v) > 0.$$

We recall the notation \mathcal{B}^n , $\tilde{t}_1(y)$, and $\tilde{t}_2(y)$: $y \in \mathcal{B}^n$ iff $p^n(y, t; \beta, \varepsilon) > 0$ for some $t \in [0, 1]$; if $y \in \mathcal{B}^n$, then $p^n(y, t; \beta, \varepsilon) > 0$ if $t \in (\tilde{t}_1(y), \tilde{t}_2(y)) \cap [0, 1]$ and $p^n(y, t; \beta, \varepsilon) = 0$ if $t \in [0, \tilde{t}_1(y)]$ or $t \in [\tilde{t}_2(y), 1]$.

We will require that for $v \in \mathbb{Z}^d$

$$(3.15) \quad \begin{aligned} u^n(y, v, t) = 0 & \text{ if } (y, t) \in (\mathcal{S}^n \setminus \mathcal{B}^n) \times [0, 1] \\ & \text{ or if } (y, t) \in \mathcal{B}^n \times ([0, \tilde{t}_1(y)] \cup [\tilde{t}_2(y), 1]). \end{aligned}$$

Actually, the values of $u^n(y, v, t)$ for such y and t are unimportant, and we assume (3.15) simply to avoid ambiguity.

Let $u^n(y, v, t)$ be a control satisfying properties (3.14) and (3.15). The controlled process is a nonstationary jump Markov process $\{\Xi^n(t), t \in [0, 1]\}$ with state space \mathcal{S}^n and, for each $y \in \mathcal{S}^n$, $v \in \mathbb{Z}^d$, and $t \in [0, 1]$, the jump intensity $u^n(y, v, t)$ from y to $y + v/n$ at time t . The generators of the controlled processes are given by the formula

$$(3.16) \quad (\mathcal{L}^n f)(y, t) \doteq n \sum_{v \in \mathbb{Z}^d} u^n(y, v, t) [f(y + v/n) - f(y)],$$

where f is any bounded real-valued function on \mathcal{S}^n . Although the process Ξ^n and the generators \mathcal{L}^n depend on the control u^n , we do not indicate this dependence in the notation.

It will turn out for the optimal control, and in fact for any control with finite cost under the cost structure to be introduced below, that $u^n(y, v, t) \rightarrow \infty$ as $t \rightarrow \tilde{t}_2(y)$ for certain values of v whenever $\tilde{t}_2(y) \leq 1$. Since we do not require that the controls be bounded for all (y, v, t) , we must take care to ensure the existence of the controlled processes. We will say that the control u^n has an associated controlled process starting at y at time t if the following hold: there exists a Markov process $\{\Xi^n(t), s \in [t, 1]\}$ on some probability space such that the process is unique in law, $\Xi^n(t) = y$ with probability 1, and

$$\varphi(\Xi^n(s), s) - \varphi(y, t) - \int_t^s \left(\frac{\partial \varphi(\Xi^n(\tau), \tau)}{\partial \tau} + \mathcal{L}^n \varphi(\Xi^n(\tau), \tau) \right) d\tau$$

is a martingale in s for $s \in [t, 1]$ for any bounded function $\varphi : \mathcal{S}^n \times [t, 1] \mapsto \mathbb{R}$ for which $\varphi(y, s)$ is continuously differentiable in s for each value of $y \in \mathcal{S}^n$. For example, if the control u^n is bounded, then such a process $\{\Xi^n(t), t \in [0, 1]\}$ exists. This leads to the following definition.

Definition 3.2. If a control u^n satisfies properties (3.14) and (3.15) and if it has an associated controlled process $\{\Xi^n(s), s \in [t, 1]\}$, then the control will be called an *admissible* control.

Cost structure. For $t \in [0, 1]$, a path $\varphi \in D([t, 1] : \mathcal{S}^n)$, a vector $\beta \in \mathbb{R}^d$, and a number $\varepsilon > 0$, we define the exit time

$$(3.17) \quad \pi(t, \varphi; \beta, \varepsilon) \doteq \inf\{s \in [t, 1] : \|\varphi(s) - s\beta\| \geq \varepsilon\}$$

and follow the convention that the infimum over the empty set is ∞ . Thus $\pi(t, \varphi; \beta, \varepsilon) = \infty$ if and only if $\|\varphi(s) - s\beta\| < \varepsilon$ for all $s \in [t, 1]$. For $a \in \mathbb{R}$ we also define the nonnegative function

$$(3.18) \quad \ell(a) \doteq \begin{cases} a \log a - a + 1 & \text{if } a \geq 0, \\ \infty & \text{if } a < 0 \end{cases}$$

and for $w \in [0, 1] \cup \{\infty\}$ the function

$$(3.19) \quad g(w) \doteq \begin{cases} 0 & \text{if } w = \infty, \\ \infty & \text{if } w \in [0, 1]. \end{cases}$$

The function ℓ has superlinear growth in the sense that

$$\lim_{a \rightarrow \infty} \ell(a)/a = \infty.$$

The definition of the stochastic control problem is completed by defining the minimal cost function

$$(3.20)$$

$$V^n(y, t; \beta, \varepsilon) \doteq \inf E_{y,t}^n \left\{ \int_t^{\pi(t, \Xi^n; \beta, \varepsilon) \wedge 1} \sum_{v \in \mathbb{Z}^d} r(\Xi^n(s), v) \ell \left(\frac{u^n(\Xi^n(s), v, s)}{r(\Xi^n(s), v)} \right) ds + g(\pi(t, \Xi^n; \beta, \varepsilon)) \right\}.$$

The infimum in this formula is taken over all admissible controls and associated controlled processes, $E_{y,t}^n\{\cdot\}$ denotes expectation conditioned on $\Xi^n(t) = y$, and when $r = u = 0$, we set $r\ell(u/r) = 0$. We need some terminology. The expected value in the definition of $V^n(y, t; \beta, \varepsilon)$ is called the *cost* with respect to the admissible control u^n . If $V^n(y, t; \beta, \varepsilon) < \infty$, then clearly the infimum in the definition of $V^n(y, t; \beta, \varepsilon)$ can be taken over all admissible controls having finite cost. In addition, the integrand depending on s which appears in the definition of $V^n(y, t; \beta, \varepsilon)$ is called the *running cost* while the term $g(\pi(t, \Xi^n; \beta, \varepsilon))$ is called the *exit cost*. For $0 \leq t_1 < t_2 \leq 1$ we refer to the term

$$E_{y,t_1}^n \left\{ \int_{t_1}^{t_2} \sum_{v \in \mathbb{Z}^d} r(\Xi^n(s), v) \ell(u^n(\Xi^n(s), v, s)/r(\Xi^n(s), v)) ds \right\}$$

as the *expected integrated running cost* over the interval $[t_1, t_2]$.

The next theorem shows that the function $q^n(y, t; \beta, \varepsilon)$ equals the minimal cost function $V^n(y, t; \beta, \varepsilon)$. We emphasize that the proof of the theorem does not require the special structure of the jump intensity function stipulated in Condition 3.1. In fact it is valid for any jump Markov process satisfying Condition 4.1, which appears in the next section and includes Condition 3.1 as a special case. Because of this, in Section 6 we will be able to apply the theorem to other jump Markov processes arising in the paper. In [6] we derive analogous representation formulas and apply them to the study of a variety of large deviation phenomena for several classes of Markov processes. Theorem 3.3 is proved in an appendix.

Theorem 3.3. *We assume Condition 3.1 or Condition 4.1. Then for each $n \in \mathbb{N}$, $y \in \mathcal{S}^n$, $t \in [0, 1]$, $\beta \in \mathbb{R}^d$, and $\varepsilon > 0$, we have*

$$q^n(y, t; \beta, \varepsilon) = V^n(y, t; \beta, \varepsilon).$$

We next state an important assumption on the processes $\{X^n(t), t \in [0, 1]\}$. The condition is stated for the original unscaled system so that it is easier to interpret.

Condition 3.4 (Communication/Controllability Condition). *Given $x \in \mathcal{S}$, let $N(x)$ be the set of points of the form $\{x + v : r(x, v) > 0\}$; thus $N(x) = x + \Sigma(x)$, where $\Sigma(x)$ denotes the support of $r(x, \cdot)$. The set $N(x)$ can be thought of as a “neighborhood” of points that are reachable from x in a single jump. We assume that there exists a positive integer K such that for each pair of points x and y in \mathcal{S} there exists a positive integer J satisfying $J \leq K\|x - y\|$ and a sequence $\{x_0, x_1, \dots, x_J\}$ such that $x_0 = x$, $x_J = y$, and $x_{j+1} \in N(x_j)$ for each $j = 0, 1, \dots, J - 1$.*

This condition can be thought of as a communication condition for the original jump Markov process and, as we will see later on, it will be used as a controllability condition for the controlled jump Markov processes $\{\Xi^n(t), t \in [0, 1], n \in \mathbb{N}\}$. The condition implies that with positive probability each point $y \in \mathcal{S}$ can be reached from any other point $x \in \mathcal{S}$ by a sequence $\{x_0, x_1, \dots, x_J\}$ that progresses through the sets $N(x_j)$ and also that the length of this sequence is bounded above by a fixed constant K times the distance from x to y .

Conditions 3.1 and 3.4 allow us to establish the following key lemma.

Lemma 3.5. *Assume Conditions 3.1 and 3.4 and let t_1 be a number in $[0, 1)$. Define the positive integer B by equation (3.3) and the positive integer K by Condition 3.4. The following conclusions hold.*

(a) *Let $\eta \in (0, 1)$ be given. Then there exist quantities $N \in \mathbb{N}$, $\delta_0 > 0$, and $\sigma \in (0, 1 - t_1)$ with the following properties. If $n \geq N$, if $\delta \in (0, \delta_0)$, and if $y \in \mathcal{S}^n$ and $\xi \in \mathcal{S}^n$ satisfy $\|y - \xi\| \leq \delta$, then there is an admissible control \bar{u}^n with the property that if Ξ^n is the associated controlled process satisfying $\Xi^n(t_1) = y$, then*

$$(3.21) \quad \mathbb{P}_{y, t_1}^n \{ \Xi^n(t_1 + \sigma) = \xi \} = 1,$$

$$(3.22) \quad \mathbb{P}_{y, t_1}^n \left\{ \sup_{s \in [t_1, t_1 + \sigma]} \|\Xi^n(s) - \xi\| \leq BK\delta \right\} = 1,$$

and the expected integrated running cost over the interval $[t_1, t_1 + \sigma]$ satisfies the inequality

$$E_{y, t_1}^n \left\{ \int_{t_1}^{t_1 + \sigma} \sum_{v \in \mathbb{Z}^d} r(\Xi^n(s), v) \ell(\bar{u}^n(\Xi^n(s), v, s) / r(\Xi^n(s), v)) ds \right\} \leq \eta.$$

For all sufficiently small $\eta \in (0, 1)$ we can take $\delta_0 = \eta / (2K)$ and $\sigma = \iota \eta$ for some positive constant ι . With this choice, $\delta_0 \rightarrow 0$ and $\sigma \rightarrow 0$ as $\eta \rightarrow 0$.

(b) *Alternatively, let $\sigma \in (0, 1 - t_1)$ be given. Then there exist quantities $N \in \mathbb{N}$, $\delta_0 > 0$, and $D \in (0, \infty)$ with the following properties. If $n \geq N$, if $\delta \in (0, \delta_0)$, and if $y \in \mathcal{S}^n$ and $\xi \in \mathcal{S}^n$ satisfy $\|y - \xi\| \leq \delta$, then there is an admissible control \bar{u}^n with the property that if Ξ^n is the associated controlled process satisfying $\Xi^n(t_1) = y$, then formulas (3.21) and (3.22) hold and*

the expected integrated running cost over the interval $[t_1, t_1 + \sigma]$ satisfies the inequality

$$E_{y, t_1}^n \left\{ \int_{t_1}^{t_1 + \sigma} \sum_{v \in \mathbb{Z}^d} r(\Xi^n(s), v) \ell(\bar{u}^n(\Xi^n(s), v, s) / r(\Xi^n(s), v)) ds \right\} \leq D.$$

Proof. The proof of part (a) is fairly lengthy. Once we have proved part (a), the proof of part (b) is almost immediate.

Proof of part (a). We give the proof for the case $t_1 = 0$. The proof for arbitrary $t_1 \in [0, 1)$ can be handled with only minor notational modifications since the problem is translation invariant in the time variable. For $n \in \mathbb{N}$ and $x \in \mathcal{S}^n$ we define the set $N^n(x) \doteq \{x + v/n : r(x, v) > 0\}$. Since we have assumed Condition 3.4, for the given y and ξ we can find a positive integer J satisfying $J \leq nK\|y - \xi\| \leq nK\delta$ and a sequence $\{y_0, y_1, \dots, y_J\}$ such that $y_0 = y$, $y_J = \xi$, and $y_{j+1} \in N^n(y_j)$ for each $j \in \{0, 1, \dots, J - 1\}$.

It is useful to introduce a second jump Markov process $\{Y^n(t), t \in [0, \infty)\}$ defined as follows. The state space of Y^n is the unbounded set $\Psi^n \doteq \{0, 1/n, 2/n, \dots\}$; $Y^n(0) = 0$; for each nonnegative integer j the jump intensity from j/n to $(j + 1)/n$ is n while the jump intensities from j/n to any other state is zero. For $\sigma > 0$ we define the probability

$$p^n(\sigma, \delta) \doteq P \{Y^n(\sigma) \geq K\delta\}$$

and the quantity

$$q^n(\sigma, \delta) \doteq -\frac{1}{n} \log p^n(\sigma, \delta).$$

Cramér’s Theorem for the sample mean of i.i.d. Poisson random variables implies that

$$(3.23) \quad \lim_{n \rightarrow \infty} q^n(\sigma, \delta) = \sigma \inf_{\{b \in \mathbb{R} : b \geq K\delta\}} \ell(b/\sigma),$$

where the function ℓ is defined in equation (3.18).

We continue with the proof of the lemma, claiming that for each $n \in \mathbb{N}$ the quantity $q^n(\sigma, \delta)$ has a representation in terms of a stochastic optimal control problem. In this case the controls are measurable functions $u^n(z, t)$ mapping $\Psi^n \times [0, 1]$ into $[0, \infty)$. Since the process $\{Y^n(t), t \in [0, \sigma]\}$ has only one type of jump, the controls do not have a v -dependence. We will say that the control u^n has an associated controlled process starting at y at time t if the following hold: there exists a Markov process $\{\Upsilon^n(t), t \in [0, \sigma]\}$ on some probability space such that the Markov process is unique in law, $\Upsilon^n(0) = 0$ with probability 1, and

$$\begin{aligned} & \varphi(\Xi^n(t), t) - \varphi(0, 0) \\ & - \int_0^t \left(\frac{\partial \varphi(\Xi^n(\tau), \tau)}{\partial \tau} + n u^n(\Upsilon^n(\tau), \tau) \right) [\varphi(\Upsilon^n(\tau) + 1/n, \tau) \\ & \qquad \qquad \qquad - \varphi(\Upsilon^n(\tau), \tau)] d\tau \end{aligned}$$

is a martingale in t for $t \in [0, \sigma]$ for any bounded function $\varphi: \mathcal{S}^n \times [0, \sigma] \rightarrow \mathbb{R}$ for which $\varphi(y, s)$ is continuously differentiable in s for each value of $y \in \Psi^n$. If a control u^n has an associated controlled process $\{\Upsilon^n(t), t \in [0, \sigma]\}$, then the control will be called an admissible control. For f any bounded real-valued function on Ψ^n and $t \in [0, \sigma]$, the associated controlled process has the generator

$$(\mathcal{A}^n f)(z, t) = n u^n(z, t) [f(z + 1/n) - f(z)].$$

The corresponding minimal cost function is given by the formula

$$V^n(\sigma, \delta) \doteq \inf E_0^n \left\{ \int_0^\sigma \ell(u^n(\Upsilon^n(s), s)) ds + g_{[0, K\delta]}(\Upsilon^n(\sigma)) \right\},$$

where the infimum is taken over all admissible controls u^n , $E_0^n\{\cdot\}$ denotes expectation conditioned on $\Upsilon^n(0) = 0$, and

$$g_{[0, K\delta]}(z) \doteq \begin{cases} 0 & \text{if } z \geq K\delta, \\ \infty & \text{if } z \in [0, K\delta). \end{cases}$$

By straightforwardly adapting the proof of Theorem 3.3, one shows that

$$q^n(\sigma, \delta) = V^n(\sigma, \delta).$$

In fact, since the present control problem does not involve an exit time, the proof of this representation is easier than the proof of Theorem 3.3.

Since the limit (3.23) holds, given any $\theta > 0$ there exists $N \in \mathbb{N}$ and for each $n \geq N$ there exists an admissible control u^n such that

$$\begin{aligned} q^n(\sigma, \delta) &\leq E_0^n \left\{ \int_0^\sigma \ell(u^n(\Upsilon^n(s), s)) ds + g_{[0, K\delta]}(\Upsilon^n(\sigma)) \right\} \\ &\leq \sigma \inf_{\{b \in \mathbb{R}: b \geq K\delta\}} \ell(b/\sigma) + \theta < \infty. \end{aligned}$$

We denote by P_0^n probability conditioned on $\Upsilon^n(0) = 0$. The last inequality implies that

$$(3.24) \quad E_0^n \{g_{[0, K\delta]}(\Upsilon^n(\sigma))\} = 0 \quad \text{or} \quad P_0^n \{\Upsilon^n(\sigma) \geq K\delta\} = 1$$

and that

$$(3.25) \quad E_0^n \left\{ \int_0^\sigma \ell(u^n(\Upsilon^n(s), s)) ds \right\} \leq \sigma \inf_{\{b \in \mathbb{R}: b \geq K\delta\}} \ell(b/\sigma) + \theta.$$

We would like to adapt the control u^n for the Υ^n process so that it can be applied to the Ξ^n process. In order to do this, it is useful to introduce an inverted tree structure with the point ξ at the top of the inverted tree. We call the top level J . The tree has the following properties: each point $y \in \mathcal{S}^n$ appears at most once in the tree; each point $y \in \mathcal{S}^n$ satisfying $\|y - \xi\| \leq \delta$ appears exactly once in the tree at level 0 or higher; if the point y appears in the tree at level $i < J$, then there is a vector v_y with $r(y, v_y) > 0$ and a point $z = y + v_y/n$ in the tree at level $i + 1$. Such a tree exists by Condition 3.4. Informally, for each $i \in \{1, 2, \dots, J\}$ level $J - i$ of the tree contains all points for which the smallest number of jumps needed to reach the point ξ is i .

We can now adapt the control u^n for the Υ^n process, which has the state space Ψ^n , so that it can be applied to the Ξ^n process, which has the state space \mathcal{S}^n . Any point $y \in \mathcal{S}^n$ that is at level j in the tree is identified with the point $j/n \in \Psi^n$. For such a point y , we define

$$\bar{u}^n(y, v, t) \doteq \begin{cases} u^n(j/n, t) & \text{if } v = v_y, \\ 0 & \text{otherwise.} \end{cases}$$

When the process Ξ^n hits the state ξ , we stop the process by defining

$$\bar{u}^n(\xi, v, t) \doteq 0 \text{ for all } v \in \mathbb{Z}^d.$$

Thus, aside from relabeling the states according to their level and stopping the process Ξ^n when it reaches ξ , the processes $\{\Upsilon^n(t), t \in [0, \sigma]\}$ and $\{\Xi^n(t), t \in [0, \sigma]\}$ are the same. For the latter process the quantity \bar{u}^n is an admissible control. Equation (3.24) and the fact that $J \leq nK\delta$ imply that

$$P_{y,0}^n\{\Xi^n(\sigma) = \xi\} = 1.$$

Let $\{y_j, j = 0, 1, \dots, \hat{J}\}$, $\hat{J} \leq J$, be a sequence of states that the process Ξ^n could possibly visit given a starting point y_0 satisfying $\|y_0 - \xi\| \leq \delta$. Then

$$\|y_j - \xi\| \leq \frac{1}{n} \sum_{i=j}^{\hat{J}-1} \|v_{y_i}\| \leq \frac{1}{n} JB \leq BK\delta.$$

Hence we obtain the equality

$$P_{y,0}^n \left\{ \sup_{s \in [0, \sigma]} \|\Xi^n(s) - \xi\| \leq BK\delta \right\} = 1.$$

We next estimate the expected integrated running cost corresponding to the use of the control \bar{u}^n . According to equation (3.4) there exists a constant $C \in [1, \infty)$ such that

$$(3.26) \quad 1/C \leq r(x, v) \leq C$$

for any $n \in \mathbb{N}$, $x \in \mathcal{S}^n$, and $v \in \mathbb{Z}^d$ satisfying $r(x, v) > 0$. For $n \in \mathbb{N}$ and $x \in \mathcal{S}^n$ let $|\Sigma(x)|$ denote the number of elements in the set $\Sigma(x)$. Condition 3.1 implies that there exists $M \in \mathbb{N}$ so that

$$(3.27) \quad \sup_{n \in \mathbb{N}} \sup_{x \in \mathcal{S}^n} |\Sigma(x)| \leq M.$$

Finally let $A \in (0, \infty)$ and $a \in (0, \infty)$ satisfy

$$r\ell(u/r) \leq A\ell(u) + a \text{ for all } u \in [0, \infty) \text{ and all } r \in [1/C, C].$$

Since $\ell(0) = 1$, we have for each $n \in \mathbb{N}$ satisfying $n \geq N$

$$\begin{aligned}
 (3.28) \quad & E_{y,0}^n \left\{ \int_0^\sigma \sum_{v \in \mathbb{Z}^d} r(\Xi^n(s), v) \ell(\bar{u}^n(\Xi^n(s), v, s)/r(\Xi^n(s), v)) ds \right\} \\
 & \leq E_0^n \left\{ \int_0^\sigma (CM + A \ell(u^n(\Upsilon^n(s), s)) + a) ds \right\} \\
 & \leq CM\sigma + a\sigma + A E_0^n \left\{ \int_0^\sigma \ell(u^n(\Upsilon^n(s), s)) ds \right\} \\
 & \leq \sigma(CM + a) + \sigma A \inf_{\{b \in \mathbb{R} : b \geq K\delta\}} \ell(b/\sigma) + A\theta.
 \end{aligned}$$

The last step uses the inequality (3.25). The conclusion of part (a) of the lemma now follows from the following choices of constants:

$$\sigma \doteq \eta/[2(CM + a)], \delta_0 \doteq \sigma/(2K), \theta \doteq \eta/2A, N \text{ picked as in the proof.}$$

Note that because $\ell(1) = 0$ the chosen value of δ_0 implies that

$$\inf_{\{b \in \mathbb{R} : b \geq K\delta\}} \ell(b/\sigma) = 0 \text{ for all } \delta \in (0, \delta_0).$$

Proof of part (b). Working with the given value of σ , we follow the proof of part (a) through display (3.28). The conclusion of part (b) holds with D given by the last line of this display. This completes the proof of Lemma 3.5. \square

Remark 3.6. The proof of Lemma 3.5 uses Condition 3.4 but not the full force of Condition 3.1. In fact, the proof uses only two consequences of Condition 3.1; namely, the bounds (3.26) and (3.27). Because of this, in Section 6 we will be able to apply the lemma to other jump Markov processes arising in the paper. \square

We introduce the set

$$\begin{aligned}
 \mathcal{S}^\infty \doteq \{ \xi \in \mathbb{R}^d : \xi = \lim_{n \rightarrow \infty} \xi^n \text{ for some sequence } \{ \xi^n, n \in \mathbb{N} \} \\
 \text{with each } \xi^n \in \mathcal{S}^n \}.
 \end{aligned}$$

Together, Conditions 3.1 and 3.4 imply that \mathcal{S}^∞ is the effective “limit state space” of the process. By this we mean the following: given any point $\xi \in \mathcal{S}^\infty$ and any $\delta > 0$, then for all sufficiently large n there exists at least one point in \mathcal{S}^n that is within δ of y and that communicates with all other points in \mathcal{S}^n . We can now prove the main result of this section.

Proposition 3.7. *Assume Conditions 3.1 and 3.4 and consider the linear subspace*

$$\Gamma \doteq \{ \beta \in \mathbb{R}^d : \langle \beta, w_i \rangle = 0 \text{ for all } i = 1, 2, \dots, \kappa \} = \mathcal{F}^{(=, \dots, =)}.$$

Let numbers $0 \leq t_1 \leq t_2 \leq 1$ and a vector $\beta \in \Gamma$ be given. For $y \in \mathcal{S}^n$ and $\varepsilon > 0$ define the conditional probability

$$P^n(y, t_1, t_2; \beta, \varepsilon) \doteq P_{y,t_1}^n \left\{ \sup_{s \in [t_1, t_2]} \|X^n(s) - (s - t_1)\beta\| < \varepsilon \right\}.$$

Then there exists a quantity $L(\beta) \in [0, \infty)$ such that the following limits hold:

$$\begin{aligned}
 & \lim_{\varepsilon \rightarrow 0} \lim_{\delta \rightarrow 0} \liminf_{n \rightarrow \infty} \inf_{\{y \in \mathcal{S}^n : \|y\| \leq \delta\}} \left(-\frac{1}{n} \log p^n(y, t_1, t_2; \beta, \varepsilon) \right) \\
 (3.29) \quad &= \lim_{\varepsilon \rightarrow 0} \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\{y \in \mathcal{S}^n : \|y\| \leq \delta\}} \left(-\frac{1}{n} \log p^n(y, t_1, t_2; \beta, \varepsilon) \right) \\
 &= (t_2 - t_1) \cdot L(\beta).
 \end{aligned}$$

In particular, taking $t_1 = 0$ and $t_2 = 1$ and using the notation $p^n(y, 0; \beta, \varepsilon) \doteq p^n(y, 0, 1; \beta, \varepsilon)$, we have

$$\begin{aligned}
 & \lim_{\varepsilon \rightarrow 0} \lim_{\delta \rightarrow 0} \liminf_{n \rightarrow \infty} \inf_{\{y \in \mathcal{S}^n : \|y\| \leq \delta\}} \left(-\frac{1}{n} \log p^n(y, 0; \beta, \varepsilon) \right) \\
 (3.30) \quad &= \lim_{\varepsilon \rightarrow 0} \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\{y \in \mathcal{S}^n : \|y\| \leq \delta\}} \left(-\frac{1}{n} \log p^n(y, 0; \beta, \varepsilon) \right) \\
 &= L(\beta).
 \end{aligned}$$

Proof. We will prove the limits in formula (3.30). The limits in formula (3.29) follows from these by a change of time variable. Concerning the limits in (3.30) we will give the proof for the case $\beta = 0$. As we remark at the end of the proof, the proof for an arbitrary vector $\beta \in \Gamma$ is almost the same.

Recall the definition (3.7)

$$q^n(y, t; \beta, \varepsilon) \doteq -\frac{1}{n} \log p^n(y, t; \beta, \varepsilon).$$

We use the representation for $q^n(y, 0; 0, \varepsilon)$ given by Theorem 3.3 and begin by first considering the case when $y = 0$. Let m and n be positive integers such that $m > n$ and define the quantity $\theta \doteq m/n$. We denote by $\sigma \in (0, 1)$ a number to be specified below and set $\Delta \doteq (1 + \sigma)/\theta$. We will partition the interval $[0, 1]$ into the intervals $[i\Delta, (i+1)\Delta)$, $i \in \{0, 1, \dots, [1/\Delta]\}$, and the interval $[\Delta[1/\Delta], 1]$. For $\alpha \in \mathbb{R}$, $[\alpha]$ denotes the integral part of α .

We write the interval $[0, \Delta)$ as

$$[0, \Delta) = [0, 1/\theta) \cup [1/\theta, (1 + \sigma)/\theta).$$

The aim of the first part of the proof is to construct for a suitable choice of $\sigma \in (0, 1)$ and for sufficiently large $m > n$ controls \hat{u}^m and \hat{u}^n on each of the subintervals $[0, 1/\theta)$ and $[0, \sigma/\theta)$ and from these controls to obtain a suitable composite control on the entire interval $[0, \Delta)$. We consider the second subinterval followed by the first one.

According to part (a) of Lemma 3.5, for any $\eta \in (0, 1)$ there exist quantities $\delta_0 > 0$, $\sigma \in (0, 1)$, and $N \in \mathbb{N}$ such that for any $n \geq N$ there exists an admissible control \bar{u}^n and an associated controlled process $\{\Xi^n(t), t \in [0, \sigma]\}$ with the following properties. If $\varepsilon \in (0, \delta_0)$ and if the process Ξ^n starts at time 0 at a point $y \in \mathcal{S}^n$ satisfying $\|y\| \leq \varepsilon$, then

- with probability 1 the control moves Ξ^n to the origin at time σ ;
- with probability 1 $\|\Xi^n(s)\| \leq BK\varepsilon$ for all $s \in [0, \sigma]$;

- the expected integrated running cost over the interval $[0, \sigma]$ is no greater than η .

For the remainder of the proof we will assume that $\varepsilon \in (0, \delta_0)$. We now adapt the control \bar{u}^n so that it can be applied to the process $\Xi^m(t)$ for t in the interval $[0, \sigma/\theta]$. We do this via the definition

$$(3.31) \quad \bar{u}^m(x, v, t) \doteq \bar{u}^n(\theta x, v, \theta t) \text{ for } x \in \mathcal{S}^m, v \in \mathbb{Z}^d, \text{ and } t \in [0, \sigma/\theta],$$

where $\theta \doteq m/n$. In the sense of equality of distributions, we have

$$(3.32) \quad \frac{1}{\theta} \Xi^n(\theta t) = \Xi^m(t) \text{ for } t \in [0, \sigma/\theta].$$

The definition (3.31) yields an admissible control \bar{u}^m with the properties that if the process Ξ^m starts at time 0 at a point $y \in \mathcal{S}^m$ satisfying $\|y\| \leq \varepsilon/\theta$, then

- with probability 1 the control moves Ξ^m to the origin at time σ/θ ;
- with probability 1 $\|\Xi^m(s)\| \leq BK\varepsilon/\theta$ for all $s \in [0, \sigma/\theta]$;
- the expected integrated running cost over the subinterval $[0, \sigma/\theta]$ is no greater than η/θ .

We now construct a suitable control \tilde{u}^m that is applied on the subinterval $[0, 1/\theta]$. Condition 3.4 and the asymptotic density property expressed in equation (3.5) imply the existence of a number $N_0 \in \mathbb{N}$ such that $q^n(0, 0; 0, \varepsilon) \in [0, \infty)$ for all $n \geq N_0$. We now take $n \geq N \vee N_0$, where the number N was defined at the beginning of the last paragraph. For $\eta \in (0, 1)$ and $n \geq N \vee N_0$ let \tilde{u}^n be an admissible control that is η -optimal in the representation for $q^n(0, 0; 0, \varepsilon)$ given in Theorem 3.3. In other words, if $\{\Xi^n(t), t \in [0, 1]\}$ is the associated controlled process, then the cost with respect to \tilde{u}^n is within η of the infimum $V^n(0, 0; 0, \varepsilon)$, or

$$E_0^n \left\{ \int_0^{\pi(0, \Xi^n; 0, \varepsilon) \wedge 1} \sum_{v \in \mathbb{Z}^d} r(\Xi^n(s), v) \ell(\tilde{u}^n(\Xi^n(s), v, s)/r(\Xi^n(s), v)) ds + g(\pi(0, \Xi^n; 0, \varepsilon)) \right\} \leq V^n(0, 0; 0, \varepsilon) + \eta = q^n(0, 0; 0, \varepsilon) + \eta$$

This implies that with probability 1 $g(\pi(0, \Xi^n; 0, \varepsilon)) = 0$, which in turn means that under the control \tilde{u}^n we have with probability 1 $\pi(0, \Xi^n; 0, \varepsilon) = 0$ and so $\|\Xi^n(s)\| < \varepsilon$ for all $s \in [0, 1]$.

We next adapt the control \tilde{u}^n so that it can be applied to the process $\Xi^m(t)$ for t in the subinterval $[0, 1/\theta]$. We do this via the analogous definition

$$(3.33) \quad \tilde{u}^m(x, v, t) \doteq \tilde{u}^n(\theta x, v, \theta t) \text{ for } x \in \mathcal{S}^m, v \in \mathbb{Z}^d, \text{ and } t \in [0, 1/\theta],$$

where $\theta \doteq m/n$. In the sense of equality of distributions, we have as in equation (3.32)

$$\frac{1}{\theta} \Xi^n(\theta t) = \Xi^m(t) \text{ for } t \in [0, 1/\theta].$$

The definition (3.33) yields an admissible control \tilde{u}^m with the properties that

- with probability 1 $\|\Xi^m(s)\| < \varepsilon/\theta$ for all $s \in [0, 1/\theta]$ and in particular $\|\Xi^m(1/\theta)\| \leq \varepsilon/\theta$;
- since $BK \geq 1$, with probability 1 $\|\Xi^m(s)\| < BK\varepsilon/\theta$ for all $s \in [0, 1/\theta]$;
- the resulting expected integrated running cost over the subinterval $[0, 1/\theta]$ is bounded above by $[q^n(0, 0; 0, \varepsilon) + \eta]/\theta$.

We now define a composite control u^m that is applied to the process $\Xi^m(t)$ for t in the interval $[0, \Delta)$. This control is given by the formula

$$u^m(x, v, t) \doteq \begin{cases} \tilde{u}^m(x, v, t) & \text{for } t \in [0, 1/\theta), \\ \tilde{u}^m(x, v, t - 1/\theta) & \text{for } t \in [1/\theta, (1 + \sigma)/\theta) = [1/\theta, \Delta). \end{cases}$$

This composite control is well defined. Indeed, the second control $\tilde{u}^m(x, v, t - 1/\theta)$ requires the process Ξ^m to start at time $1/\theta$ at a point $y \in \mathcal{S}^m$ satisfying $\|y\| \leq \varepsilon/\theta$, and under the first control $\tilde{u}^m(x, v, t)$ the process Ξ^m at time $1/\theta$ satisfies $\|\Xi^m(1/\theta)\| < \varepsilon/\theta$ with probability 1. Under the use of the composite control, with probability 1 the process Ξ^m starts at the origin at time 0 and returns to the origin at time Δ . In addition, when this control is used, the expected integrated running cost over the interval $[0, \Delta)$ is no greater than $[q^n(0, 0; 0, \varepsilon) + 2\eta]/\theta$. Since with probability 1 the composite control u^m returns the process to the origin at time Δ , suitable translations of the same control can be used for t in each interval of the form $[i\Delta, i\Delta + \Delta)$, $i \in \{0, 1, \dots, [1/\Delta]\}$, and for t in the last interval $[\Delta[1/\Delta], 1]$. We denote the composite control over the entire interval $[0, 1]$ by the same symbol u^m . Under the control u^m the expected integrated running cost over the interval $[0, 1]$ satisfies the inequality

$$\begin{aligned} E_{y,0}^m \left\{ \int_0^1 \sum_{v \in \mathbb{Z}^d} r(\Xi^m(s), v) \ell(u^m(\Xi^m(s), v, s)/r(\Xi^m(s), v)) ds \right\} \\ \leq \frac{1 + \theta}{\theta} [q^n(0, 0; 0, \varepsilon) + 2\eta]. \end{aligned}$$

Under the control u^m we also have $\|\Xi^m(s)\| < BK\varepsilon/\theta$ for all $s \in [0, 1]$ with probability 1. Thus if $n \geq N \vee N_0$ and $m > BK n$, then $\|\Xi^m(s)\| < \varepsilon$ for all $s \in [0, 1]$ with probability 1, and so with probability 1 the exit cost $g(\pi(0, \Xi^m; 0, \varepsilon))$ equals 0. The representation formula $q^m(0, 0; 0, \varepsilon) = V^m(0, 0; 0, \varepsilon)$ (Theorem 3.3), the definition of $V^m(0, 0; 0, \varepsilon)$ as an infimum over admissible controls, and the bound on the expected integrated running cost

given in the last display now yield

$$\begin{aligned}
 & q^m(0, 0; 0, \varepsilon) \\
 (3.34) \quad & \leq E_{y,0}^m \left\{ \int_0^1 \sum_{v \in \mathbb{Z}^d} r(\Xi^m(s), v) \ell(u^m(\Xi^m(s), v, s)/r(\Xi^m(s), v)) ds \right\} \\
 & \leq \frac{1 + \theta}{\theta} [q^n(0, 0; 0, \varepsilon) + 2\eta].
 \end{aligned}$$

This is the type of bound that arises in proving the existence of limits involving subadditive functions (see [13, Lemma 3.11]).

We are now ready to prove the limits in formula (3.30) for $\beta = 0$. For each $n \geq N \vee N_0$ the quantities $q^n(0, 0; 0, \varepsilon)$ are monotonically nondecreasing as $\varepsilon \rightarrow 0$. Hence the quantity

$$L(0) \doteq \lim_{\varepsilon \rightarrow 0} \liminf_{n \rightarrow \infty} q^n(0, 0; 0, \varepsilon)$$

is well defined. We prove below that $L(0)$ is finite. In the inequality (3.34) we take first $m \rightarrow \infty$ and then $n \rightarrow \infty$. We find that

$$\limsup_{m \rightarrow \infty} q^m(0, 0; 0, \varepsilon) \leq L(0) + 2\eta$$

and therefore that

$$\lim_{\varepsilon \rightarrow 0} \limsup_{m \rightarrow \infty} q^m(0, 0; 0, \varepsilon) \leq L(0) + 2\eta.$$

Sending $\eta \rightarrow 0$, we obtain

$$\lim_{\varepsilon \rightarrow 0} \limsup_{m \rightarrow \infty} q^m(0, 0; 0, \varepsilon) \leq L(0).$$

These calculations show that

$$(3.35) \quad \lim_{\varepsilon \rightarrow 0} \limsup_{m \rightarrow \infty} q^m(0, 0; 0, \varepsilon) = \lim_{\varepsilon \rightarrow 0} \liminf_{m \rightarrow \infty} q^m(0, 0; 0, \varepsilon) = L(0).$$

The desired limits in formula (3.30) involve the infimum and the supremum over the set $\{y \in \mathcal{S}^n : \|y\| \leq \delta\}$, where δ is sent to zero before ε is sent to zero. Fix $\varepsilon \in (0, 1)$ and let $\eta \in (0, 1)$ be given. Using the representation formula in Theorem 3.3 and part (a) of Lemma 3.5, we see that there exist $N \in \mathbb{N}$, $\delta_0 > 0$, and $\sigma \in (0, 1)$ such that if $n \geq N$, if $\delta \in (0, \delta_0)$, and if $y \in \mathcal{S}^n$ satisfies $\|y\| \leq \delta$, then

$$|q^n(y, 0; 0, \varepsilon) - q^n(0, \sigma; 0, \varepsilon)| = |V^n(y, 0; 0, \varepsilon) - V^n(0, \sigma; 0, \varepsilon)| \leq \eta.$$

Furthermore, for all sufficiently small $\eta \in (0, 1)$ we can take $\sigma = \iota\eta$ for some constant ι . We have already proved the limit in equation (3.35), and by a change of time variable we can obtain from this the limit

$$\lim_{\varepsilon \rightarrow 0} \liminf_{m \rightarrow \infty} q^m(0, \sigma; 0, \varepsilon) = \lim_{\varepsilon \rightarrow 0} \limsup_{m \rightarrow \infty} q^m(0, \sigma; 0, \varepsilon) = (1 - \sigma) \cdot L(0).$$

Hence there exists $N_1 \in \mathbb{N}$ such that for any $n \geq N \vee N_1$

$$|q^n(0, 0; 0, \varepsilon) - q^n(0, \sigma; 0, \varepsilon)| \leq (2 + L(0)\iota)\eta.$$

It follows that for any $n \geq N \vee N_1$, $\delta \in (0, \delta_0)$, $y \in \mathcal{S}^n$ satisfying $\|y\| \leq \delta$, and $\varepsilon \in (0, 1)$

$$|q^n(y, 0; 0, \varepsilon) - q(0, 0; 0, \varepsilon)| \leq (3 + L(0)t)\eta.$$

The limits in formula (3.30) for $\beta = 0$ follow from this bound; from the finiteness of $L(0)$, which is proven next; and from the limit (3.35).

We now prove that $L(0) \in [0, \infty)$. Let $\varepsilon \in (0, 1)$ be given and let y be any point in \mathcal{S}^n that satisfies $\|y\| \leq \varepsilon/2BK$. According to part (b) of Lemma 3.5 with $t_1 \doteq 0$ and $\sigma \doteq 1$, for all sufficiently large $n \in \mathbb{N}$ and all sufficiently small $\varepsilon \in (0, 1)$ there exists a quantity $D \in (0, \infty)$ that is independent of ε and an admissible control u^n with the following properties. If $\{\Xi^n(t), t \in [0, 1]\}$ is the associated controlled process and if $\Xi^n(0) = 0$, then with probability 1 $\Xi^n(1) = y$, $\sup_{s \in [0, 1]} \|\Xi^n(s)\| \leq \varepsilon/2$, and the expected integrated running cost over the interval $[0, 1]$ corresponding to the use of the control u^n is bounded above by D . The second of these properties implies that with probability 1 the exit cost $g(\pi(0, \Xi^n; 0, \varepsilon))$ equals 0. The representation formula $q^n(0, 0; 0, \varepsilon) = V^n(0, 0; 0, \varepsilon)$ (Theorem 3.3), the definition of $V^n(0, 0; 0, \varepsilon)$ as an infimum over admissible controls, and the bound on the expected integrated running cost just given imply that for all sufficiently small $\varepsilon \in (0, 1)$ and all sufficiently large $n \in \mathbb{N}$

$$q^n(0, 0; 0, \varepsilon) \leq D < \infty$$

and thus that $L(0) \in [0, D]$.

For $\beta = 0$ we have completed the proof of the limits in formula (3.30) and thus the proof of the limits in formula (3.29). In order to prove the limits for arbitrary $\beta \in \Gamma$ we proceed in essentially the same way. We recall in the proof just given for the case $\beta = 0$ that the control returns the process $\Xi^n(t)$ to the origin at each time $t = i\Delta$, $i \in \{1, \dots, \lfloor 1/\Delta \rfloor\}$. In order to treat the case of arbitrary $\beta \in \Gamma$ the control will be chosen to return the process at each of these times to a suitable point that is in \mathcal{S}^m and that is near the point $i\Delta\beta$. \square

The next proposition gives a convexity property of the function $L(\cdot)$.

Proposition 3.8. *Assume Conditions 3.1 and 3.4 and consider the linear subspace*

$$\Gamma \doteq \{\beta \in \mathbb{R}^d : \langle \beta, w_i \rangle = 0 \text{ for all } i = 1, 2, \dots, \kappa\} = \mathcal{F}^{(=, \dots, =)}.$$

Let L be the function defined on Γ that satisfies the limits in equation (3.29). Then L is a finite convex function on Γ .

Proof. Let $\rho^{(1)}, \rho^{(2)} \in \mathbb{R}$ and $\beta^{(1)}, \beta^{(2)} \in \Gamma$ satisfy $\rho^{(1)} > 0$, $\rho^{(2)} > 0$, and $\rho^{(1)} + \rho^{(2)} = 1$. Define the quantity $\beta \doteq \rho^{(1)}\beta^{(1)} + \rho^{(2)}\beta^{(2)}$. According to Proposition 3.7, for any $0 \leq t_1 \leq t_2 \leq 1$ and each $i = 1, 2$

$$\begin{aligned} & \lim_{\varepsilon \rightarrow 0} \lim_{\delta \rightarrow 0} \liminf_{n \rightarrow \infty} \inf_{\{y \in \mathcal{S}^n : \|y\| \leq \delta\}} \left(-\frac{1}{n} \log p^n(y, t_1, t_2; \beta^{(i)}, \varepsilon) \right) \\ (3.36) \quad & = \lim_{\varepsilon \rightarrow 0} \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\{y \in \mathcal{S}^n : \|y\| \leq \delta\}} \left(-\frac{1}{n} \log p^n(y, t_1, t_2; \beta^{(i)}, \varepsilon) \right) \\ & = (t_2 - t_1) \cdot L(\beta^{(i)}). \end{aligned}$$

Let $\varepsilon > 0$ be given and choose $N \in \mathbb{N}$ such that

$$N^{-1} \max\{\rho^{(1)}\|\beta^{(1)} - \beta\|, \rho^{(2)}\|\beta^{(2)} - \beta\|\} \leq \varepsilon/2.$$

Define $\delta_{-1}^{(2)} = 0$ and choose numbers $\delta_0^{(1)}, \delta_0^{(2)}, \delta_1^{(1)}, \delta_1^{(2)}, \dots, \delta_{N-1}^{(1)}, \delta_{N-1}^{(2)}$ satisfying the inequalities

$$0 < \delta_0^{(1)} \leq \delta_0^{(2)} \leq \delta_1^{(1)} \leq \delta_1^{(2)} \leq \dots \leq \delta_{N-1}^{(1)} \leq \delta_{N-1}^{(2)} \leq \varepsilon/4.$$

It follows from the Markov property and the fact that the probabilities are unchanged by translations of the initial condition in Γ that

$$\begin{aligned} & -\frac{1}{n} \log p^n(0, 0; \beta, \varepsilon) \\ & \leq \sum_{j=0}^{N-1} \sup_{\{y \in \mathcal{S}^n : \|y\| \leq \delta_{j-1}^{(2)}\}} \left(-\frac{1}{n} \log p^n(y, j/N, (j + \rho^{(1)})/N; \beta^{(1)}, \delta_j^{(1)}) \right) \\ & \quad + \sum_{j=0}^{N-1} \sup_{\{y \in \mathcal{S}^n : \|y\| \leq \delta_j^{(1)}\}} \left(-\frac{1}{n} \log p^n(y, (j + \rho^{(1)})/N, (j + 1)/N; \beta^{(2)}, \delta_j^{(2)}) \right). \end{aligned}$$

We take limits in this inequality in the following order: $n \rightarrow \infty$, $\delta_0^{(1)} \rightarrow 0$, $\delta_0^{(2)} \rightarrow 0$, $\delta_1^{(1)} \rightarrow 0$, \dots , $\delta_{N-1}^{(1)} \rightarrow 0$, $\delta_{N-1}^{(2)} \rightarrow 0$, $\varepsilon \rightarrow 0$. By using formula (3.36), we obtain

$$L(\beta) = L(\rho^{(1)}\beta^{(1)} + \rho^{(2)}\beta^{(2)}) \leq \rho^{(1)}L(\beta^{(1)}) + \rho^{(2)}L(\beta^{(2)}).$$

This proves that $L(\cdot)$ is convex on Γ . The finiteness of $L(\cdot)$ was shown in Proposition 3.7. \square

The next lemma gives an inequality that will be used in the last result in this section. We omit the straightforward proof, which is based on elementary inequalities.

Lemma 3.9. *Let $\gamma > 0$ and $C \in [1, \infty)$ be given. Then there exists $\theta \in (0, \infty)$ such that for all $u \in [0, \infty)$ and real numbers r_1 and r_2 satisfying*

$$|r_1 - r_2| \leq \theta, \quad 1/C \leq r_1 \leq C, \quad 1/C \leq r_2 \leq C,$$

we have the inequality

$$(1 - \gamma)r_2\ell(u/r_2) - \gamma \leq r_1\ell(u/r_1) \leq (1 + \gamma)r_2\ell(u/r_2) + \gamma.$$

The nonnegative function ℓ is defined in formula (3.18).

For use in Section 6 we want to consider a parametrized family of jump intensity functions $r_\lambda(x, v)$, where λ takes values in some subset Λ of \mathbb{R}^d . For a fixed $\lambda_0 \in \Lambda$ and each point $x \in \mathcal{S}$ we define the set

$$\Sigma_0(x) \doteq \{v \in \mathbb{Z}^d : r_{\lambda_0}(x, v) > 0\}.$$

We have the following fact.

Lemma 3.10. *Suppose that in place of the jump intensity function $r(x, v)$ we consider a parameterized family of jump intensity functions $r_\lambda(x, v)$, where λ*

takes values in some subset Λ of \mathbb{R}^d . Assume that Conditions 3.1 and 3.4 hold with $r(\cdot, \cdot)$ replaced by $r_\lambda(\cdot, \cdot)$ for each $\lambda \in \Lambda$, that $r_\lambda(x, v) = 0$ for each $x \in \mathcal{S}$ and $v \in \mathbb{Z}^d \setminus \Sigma_0(x)$, that the function mapping $\lambda \in \Lambda \mapsto r_\lambda(x, v)$ is uniformly continuous for each $x \in \mathcal{S}$ and $v \in \Sigma_0(x)$, and that there exists $C \in [1, \infty)$ such that for all $x \in \mathcal{S}$ and $v \in \Sigma_0(x)$

$$(3.37) \quad 1/C \leq \inf_{\lambda \in \Lambda} r_\lambda(x, v) \leq \sup_{\lambda \in \Lambda} r_\lambda(x, v) \leq C.$$

Finally, for each $\lambda \in \Lambda$ and $\beta \in \Gamma$ we define the quantity $L_\lambda(\beta)$ by Proposition 3.7 with $r(x, v)$ replaced by $r_\lambda(x, v)$. Then for each $\beta \in \Gamma$ the function mapping $\lambda \in \Lambda \mapsto L_\lambda(\beta)$ is uniformly continuous.

Proof. The inequality (3.37) implies that for each point $x \in \mathcal{S}$ and $\lambda \in \Lambda$

$$\{v \in \mathbb{Z}^d : r_\lambda(x, v) > 0\} = \{v \in \mathbb{Z}^d : r_{\lambda_0}(x, v) > 0\} = \Sigma_0(x).$$

For $n \in \mathbb{N}$, $\lambda \in \Lambda$, $y \in \mathcal{S}^n$, $t \in [0, 1]$, $\beta \in \mathbb{R}^d$, and $\varepsilon \in (0, 1)$, we denote by $q_\lambda^n(y, t; \beta, \varepsilon)$ and $V_\lambda^n(y, t; \beta, \varepsilon)$ the quantities obtained from $q^n(y, t; \beta, \varepsilon)$ and $V^n(y, t; \beta, \varepsilon)$ by replacing the original jump intensity function $r(\cdot, \cdot)$ by $r_\lambda(\cdot, \cdot)$. Theorem 3.3 yields the representation formula $q_\lambda^n(y, t; \beta, \varepsilon) = V_\lambda^n(y, t; \beta, \varepsilon)$. The present lemma is a direct consequence of this equality, the definition of $L_\lambda(\beta)$, the fact that $\Sigma_0(x)$ is a finite set for each $x \in \mathcal{S}$, and Lemma 3.9. \square

This completes our study of the local model. In the next section we formulate the large deviation principle for a class of processes that extends the one just considered.

4. STATEMENT OF THE LARGE DEVIATION PRINCIPLE

In Theorem 4.3 we state the large deviation principle for a class of processes that extends the class considered in the last section. The point of the extension is to allow some additional state dependencies in the jump intensities. The proof of the theorem involves a localization argument that will allow us to apply the results of the previous section. All of the notations and conventions of Section 3 remain in effect except for part (b) of Condition 3.1, in which we assumed that the jump intensities are constant on each facet $\mathcal{F}^{(o_1, \dots, o_k)} \cap \mathcal{S}$ of \mathcal{S} . By using additional localization methods we can extend the range of applicability of Theorem 4.3 to include a number of other queueing systems, such as systems with finite buffers. An example of how this is done is given in Example 4.4 at the end of the section.

In order to define the processes of the present section, we are given a state space \mathcal{S} which is a subset of \mathbb{Z}^d containing the origin. For each $n \in \mathbb{N}$ we define the scaled state space

$$\mathcal{S}^n \doteq \left\{ \frac{1}{n} y : y \in \mathcal{S} \right\},$$

and the limit state space

$$\mathcal{S}^\infty \doteq \{ \xi \in \mathbb{R}^d : \xi = \lim_{n \rightarrow \infty} \xi^n \text{ for some sequence } \{ \xi^n, n \in \mathbb{N} \}$$

with each $\xi^n \in \mathcal{S}^n$.

We are also given a jump intensity function $r(\cdot, \cdot)$ that maps $\mathbb{R}^d \times \mathbb{Z}^d$ into $[0, \infty)$. For $x \in \mathcal{S}$ and $v \in \mathbb{Z}^d$ $r(x, v) > 0$ only if $x + v \in \mathcal{S}$. For each $x \in \mathbb{R}^d$ the support of $r(x, \cdot)$ is defined to be the set

$$\Sigma(x) \doteq \{v \in \mathbb{Z}^d : r(x, v) > 0\}.$$

Condition 3.1 is now replaced by the following condition, which generalizes it. Parts (a) and (b) have already been discussed immediately after Condition 3.1.

Condition 4.1. *We consider a finite collection of nonzero vectors $\{w^{(i)}, i = 1, 2, \dots, \kappa\}$ in \mathbb{R}^d such that each $w^{(i)}$ has rational components. We define the sets*

$$\mathcal{F}^{(\circ_1, \dots, \circ_\kappa)} \doteq \{y \in \mathbb{R}^d : \langle y, w^{(i)} \rangle \circ_i 0 \text{ for each } i = 1, 2, \dots, \kappa\},$$

where for each $i \in \{1, 2, \dots, \kappa\}$ \circ_i represents one of the three symbols $>$, $=$, $<$. The subsets $\{\mathcal{F}^{(\circ_1, \dots, \circ_\kappa)}\}$ and $\{\mathcal{F}^{(\circ_1, \dots, \circ_\kappa)} \cap \mathcal{S}\}$ are partitions of \mathbb{R}^d and of \mathcal{S} , respectively. We refer to these sets as the facets of \mathbb{R}^d and as the facets of \mathcal{S} , respectively. For each sequence of symbols $(\circ_1, \dots, \circ_\kappa)$ we assume the following:

- (a) If $\mathcal{F}^{(\circ_1, \dots, \circ_\kappa)} \cap \mathcal{S} \neq \emptyset$, then $\mathcal{F}^{(\circ_1, \dots, \circ_\kappa)} \cap \mathbb{Z}^d \subset \mathcal{S}$.
- (b) The support $\Sigma(x)$ is independent of $x \in \mathcal{F}^{(\circ_1, \dots, \circ_\kappa)} \cap \mathcal{S}$ and consists of a finite number of vectors.
- (c) There exists a continuous function $r^{(\circ_1, \dots, \circ_\kappa)}$ mapping $\mathbb{R}^d \times \mathbb{Z}^d$ into \mathbb{R} such that

$$0 < \inf_{x \in \mathbb{R}^d, v \in \Sigma(x)} r^{(\circ_1, \dots, \circ_\kappa)}(x, v) \leq \sup_{x \in \mathbb{R}^d, v \in \Sigma(x)} r^{(\circ_1, \dots, \circ_\kappa)}(x, v) < \infty.$$

We then assume that $r(\cdot, \cdot)$ has the form

$$r(x, v) = r^{(\circ_1, \dots, \circ_\kappa)}(x, v) \text{ for } x \in \mathcal{F}^{(\circ_1, \dots, \circ_\kappa)} \text{ and } v \in \mathbb{Z}^d.$$

Note that with this definition of $r(x, v)$ we automatically obtain the uniform continuity of $r(x, v)$ on the intersection of any facet $\mathcal{F}^{(\circ_1, \dots, \circ_\kappa)}$ with any bounded set. Condition 4.1 implies the existence of a constant $C \in [1, \infty)$ such that

$$(4.1) \quad 1/C \leq r(x, v) \leq C$$

for any $n \in \mathbb{N}$, $x \in \mathcal{S}^n$, and $v \in \mathbb{Z}^d$ satisfying $r(x, v) > 0$. For $n \in \mathbb{N}$ and $x \in \mathcal{S}^n$ let $|\Sigma(x)|$ denote the number of elements in the set $\Sigma(x)$. Condition 4.1 also implies that there exists $M \in \mathbb{N}$ so that

$$(4.2) \quad \sup_{n \in \mathbb{N}} \sup_{x \in \mathcal{S}^n} |\Sigma(x)| \leq M.$$

Part (c) of this condition can be weakened, and it is sometimes useful to do so. Throughout this section we will also assume the communication/controllability condition expressed in Condition 3.4.

Because of the additional state dependence that is allowed in this section, in general the scaled processes cannot simply be defined to be $X(nt)/n$ for some unscaled process $X(t)$ as in the previous section. Instead, for each $n \in \mathbb{N}$ we consider a stationary jump Markov process $\{X^n(t), t \in [0, 1]\}$ with state space

\mathcal{S}^n and generator

$$\mathcal{L}^n f(y) \doteq n \sum_{v \in \mathbb{Z}^d} r(y, v) [f(y + v/n) - f(y)],$$

where f is any bounded real-valued function on \mathcal{S}^n . The quantity $r(y, v)$ gives the jump intensity from y to $y + v/n$. Although the processes $X^n(t)$ and $X(nt)/n$ coincided in the previous section, they in general will differ here. As in Section 3, for $n \in \mathbb{N}$, $y \in \mathbb{R}^d$, $t \in [0, 1]$, $\beta \in \mathbb{R}^d$, and $\varepsilon > 0$ we define the functions

$$(4.3) \quad \begin{aligned} q^n(y, t; \beta, \varepsilon) &\doteq -\frac{1}{n} \log p^n(y, t; \beta, \varepsilon) \\ &\doteq -\frac{1}{n} \log P_{y,t}^n \left\{ \sup_{s \in [t, 1]} \|X^n(s) - s\beta\| < \varepsilon \right\}. \end{aligned}$$

By Theorem 3.3 $q^n(y, t; \beta, \varepsilon)$ equals the minimal cost function $V^n(y, t; \beta, \varepsilon)$ defined in equation (3.20).

In order to formulate and prove the large deviation principle for the processes $\{X^n(t), t \in [0, 1], n \in \mathbb{N}\}$, we use the fact that large deviation behavior is “local.” By this we mean that if the starting point is x , then the large deviation behavior over short time intervals depends only on the jump intensities and the form of their discontinuities near x . Thus it makes sense to “localize” the jump intensity function around each point $x \in \mathcal{S}^\infty$, obtaining processes like those studied in Section 3. This requires the description of those facets of \mathbb{R}^d needed to study the process in the neighborhood of x . In order to do this we need some new notation, which is somewhat complicated but unavoidably so.

Let $r(x, v)$ be a jump intensity function that satisfies Condition 4.1. Thus $r(x, v)$ is continuous, and not necessarily constant, on each facet of \mathbb{R}^d . The idea is to localize, for each $\tilde{x} \in \mathcal{S}^\infty$, such a jump intensity function, obtaining a jump intensity function $\tilde{r}(\tilde{x}; y, v)$ that is constant on a new set of facets of \mathbb{R}^d and thus satisfies Condition 3.1. This will allow us to define a “local” version of the jump Markov processes $\{X^n(t), t \in [0, 1], n \in \mathbb{N}\}$ with jump intensity function $\tilde{r}(\tilde{x}; y, v)$, obtaining a sequence of jump Markov processes to which the results of Section 3 can be applied.

The queueing model is defined in terms of a fixed set of nonzero vectors $\{w^{(i)}, i = 1, 2, \dots, \kappa\}$ as described in Condition 4.1. For each point $\tilde{x} \in \mathcal{S}^\infty$ we define the natural subset of $\{1, 2, \dots, \kappa\}$ that is associated with this point by the formula

$$\Phi(\tilde{x}) \doteq \{i \in \{1, 2, \dots, \kappa\} : \langle w^{(i)}, \tilde{x} \rangle = 0\}.$$

If $\Phi(\tilde{x}) \neq \emptyset$, then we define the linear subspace

$$\Gamma(\tilde{x}) \doteq \{\beta \in \mathbb{R}^d : \langle \beta, w^{(i)} \rangle = 0 \text{ for all } i \in \Phi(\tilde{x})\}.$$

If $\Phi(\tilde{x}) = \emptyset$, then $\Gamma(\tilde{x})$ is defined to be \mathbb{R}^d . In general, $\Gamma(\tilde{x})$ is the smallest linear subspace of \mathbb{R}^d containing the facet in which \tilde{x} lies.

For each $\tilde{x} \in \mathcal{S}^\infty$ let $\kappa(\tilde{x})$ be the number of elements in $\Phi(\tilde{x})$. If $\kappa(\tilde{x}) = 0$, then $\Phi(\tilde{x}) = \emptyset$. If $\kappa(\tilde{x}) \geq 1$, then by relabeling, if necessary, the vectors $\{w^{(i)}, i = 1, 2, \dots, \kappa\}$, we can assume without loss of generality that $\Phi(\tilde{x}) =$

$\{1, 2, \dots, \kappa(\tilde{x})\}$. When $\kappa(\tilde{x}) \geq 1$, we define the new facets

$$(4.4) \quad \tilde{\mathcal{F}}^{(\circ_1, \dots, \circ_{\kappa(\tilde{x})})} \doteq \{y \in \mathbb{R}^d : \langle y - \tilde{x}, w^{(i)} \rangle \circ_i 0 \text{ for all } i = 1, \dots, \kappa(\tilde{x})\},$$

where each \circ_i represents one of the symbols $>, =, <$. If $\kappa(\tilde{x}) = 0$, then we define $\tilde{\mathcal{F}}^{(\circ_1, \dots, \circ_{\kappa(\tilde{x})})} \doteq \mathbb{R}^d$. In this notation $\Gamma(\tilde{x})$ is the facet $\tilde{\mathcal{F}}^{(=, \dots, =)}$. We call each new facet $\tilde{\mathcal{F}}^{(\circ_1, \dots, \circ_{\kappa(\tilde{x})})}$ a facet determined by \tilde{x} . The facets $\{\tilde{\mathcal{F}}^{(\circ_1, \dots, \circ_{\kappa(\tilde{x})})}\}$ determined by \tilde{x} give a partition of \mathbb{R}^d .

For $\tilde{x} \in \mathcal{S}^\infty$, y lying in a facet $\tilde{\mathcal{F}}^{(\circ_1, \dots, \circ_{\kappa(\tilde{x})})}$, and $v \in \mathbb{Z}^d$ the “local” jump intensity function $\tilde{r}(\tilde{x}; y, v)$ is defined by the formula

$$(4.5) \quad \tilde{r}(\tilde{x}; y, v) = \lim r(z, v).$$

The limit is taken as $z \rightarrow \tilde{x}$ with $z - \tilde{x}$ lying in the same facet $\tilde{\mathcal{F}}^{(\circ_1, \dots, \circ_{\kappa(\tilde{x})})}$ as y . By part (c) of Condition 4.1 the local jump intensity function is well defined. It is constant on each of the facets $\tilde{\mathcal{F}}^{(\circ_1, \dots, \circ_{\kappa(\tilde{x})})}$ determined by \tilde{x} and so satisfies Condition 3.1 with state space depending on \tilde{x} . We denote this \tilde{x} -dependent state space by $\mathcal{S}_{\tilde{x}}$. $\mathcal{S}_{\tilde{x}}$ equals the union of the sets $\tilde{\mathcal{F}}^{(\circ_1, \dots, \circ_{\kappa(\tilde{x})})} \cap \mathbb{Z}^d$, where the union is taken over those facets determined by \tilde{x} which have nonempty intersection with the original state space \mathcal{S} .

For each $\tilde{x} \in \mathcal{S}^\infty$ the introduction of two sets of facets—namely, the original facets $\mathcal{F}^{(\circ_1, \dots, \circ_\kappa)}$ and the facets $\tilde{\mathcal{F}}^{(\circ_1, \dots, \circ_{\kappa(\tilde{x})})}$ determined by \tilde{x} —may seem unnecessarily confusing. However, the introduction of the latter facets is essential for the statement of the large deviation principle in Theorem 4.3. Without them we cannot define the integrand $L(\cdot, \cdot)$ appearing in formula (4.8), which in turn gives the definition of the rate function in the theorem.

An original facet $\mathcal{F}^{(\circ_1, \dots, \circ_\kappa)}$ is said to be adjacent to \tilde{x} if the facet does not contain \tilde{x} but the closure of the facet contains \tilde{x} . For each $\tilde{x} \in \mathcal{S}^\infty$ there is a basic geometric relationship between the facets determined by \tilde{x} and a subset of the original facets. Namely, an original facet $\mathcal{F}^{(\circ_1, \dots, \circ_\kappa)}$ either contains \tilde{x} or is adjacent to \tilde{x} if and only if $\mathcal{F}^{(\circ_1, \dots, \circ_\kappa)}$ is a subset of a facet $\tilde{\mathcal{F}}^{(\circ_1, \dots, \circ_{\kappa(\tilde{x})})}$ determined by \tilde{x} . This establishes a one-to-one correspondence between the set of all facets $\tilde{\mathcal{F}}^{(\circ_1, \dots, \circ_{\kappa(\tilde{x})})}$ determined by \tilde{x} and the subset of original facets $\mathcal{F}^{(\circ_1, \dots, \circ_\kappa)}$ that either contain \tilde{x} or are adjacent to \tilde{x} . Each of the facets determined by \tilde{x} is obtained from the corresponding original facet by taking away an appropriate “distant” boundary. The definition (4.5) of $\tilde{r}(\tilde{x}; y, v)$ involves values of $r(z, v)$ only for z in facets $\mathcal{F}^{(\circ_1, \dots, \circ_\kappa)}$ that either contain \tilde{x} or are adjacent to \tilde{x} .

In Figure 3 we elucidate, in the case of the modified join-the-shorter-queue model of Section 2, the construction of the new facets which has just been discussed. We consider the example of a point \tilde{x} lying in the halfline $\mathcal{F}^{(>, >, =)}$, for which the new facets are

$$\begin{aligned} \tilde{\mathcal{F}}^{(=)} &\doteq \{y \in \mathbb{R}^2 : y_1 = y_2\}, & \tilde{\mathcal{F}}^{(>)} &\doteq \{y \in \mathbb{R}^2 : y_1 > y_2\}, \\ \tilde{\mathcal{F}}^{(<)} &\doteq \{y \in \mathbb{R}^2 : y_1 < y_2\}. \end{aligned}$$

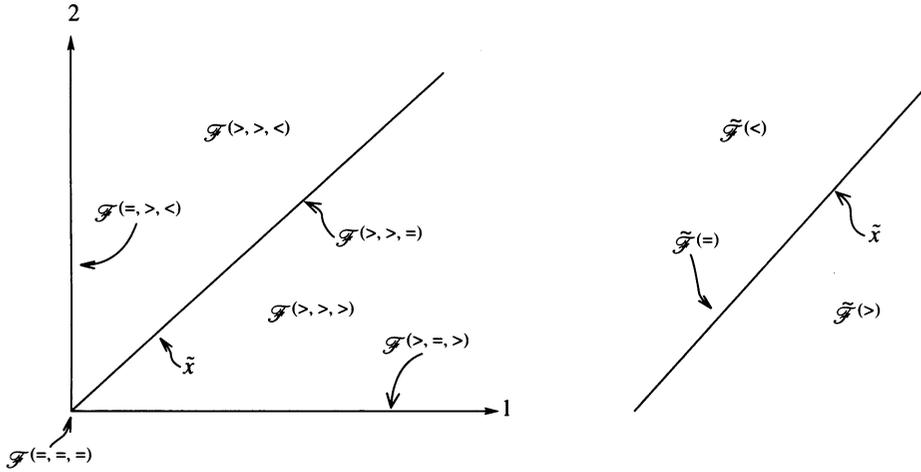


FIGURE 3. Facets of the modified join-the-shorter-queue model.

We return to the general case. Fix $\tilde{x} \in \mathcal{S}^\infty$ and for $n \in \mathbb{N}$ consider the scaled state space

$$\mathcal{S}_{\tilde{x}}^n \doteq \left\{ \frac{1}{n} y : y \in \mathcal{S}_{\tilde{x}} \right\}.$$

We define $\{\tilde{X}_{\tilde{x}}(t), t \in [0, 1]\}$ to be the stationary jump Markov process having the state space $\mathcal{S}_{\tilde{x}}$ and the jump intensity function $\tilde{r}(\tilde{x}; y, v)$. In terms of this process we define the scaled jump Markov process $\{\tilde{X}_{\tilde{x}}^n(t), t \in [0, 1]\}$ by the formula

$$\tilde{X}_{\tilde{x}}^n(t) \doteq \frac{1}{n} \tilde{X}_{\tilde{x}}(nt).$$

The state space of this scaled process is $\mathcal{S}_{\tilde{x}}^n$. For f any bounded real-valued function on $\mathcal{S}_{\tilde{x}}^n$, the generator of the scaled process is given by the formula

$$\tilde{\mathcal{L}}_{\tilde{x}}^n f(\xi) = n \sum_{v \in \mathbb{Z}^d} \tilde{r}(\tilde{x}; \xi, v) [f(\xi + v/n) - f(\xi)].$$

These processes will play an important role in the proof of the large deviation principle for the original jump Markov processes $\{X^n(t), t \in [0, 1], n \in \mathbb{N}\}$. This large deviation principle will be stated in Theorem 4.3 at the end of this section. The processes $\{\tilde{X}_{\tilde{x}}^n(t), t \in [0, \Delta], n \in \mathbb{N}\}$ provide a good approximation to the processes $\{X^n(t), t \in [0, \Delta], n \in \mathbb{N}\}$ when $\Delta > 0$ is sufficiently small.

Both Conditions 3.1 and 3.4 of Section 3 are met for the processes $\{\tilde{X}_{\tilde{x}}^n(t), t \in [0, 1], n \in \mathbb{N}\}$. For $y \in \mathcal{S}_{\tilde{x}}^n$, numbers t_1 and t_2 satisfying $0 \leq t_1 \leq t_2 \leq 1$, $\beta \in \mathbb{Z}^d$, and $\varepsilon > 0$, let $\tilde{p}_{\tilde{x}}^n(y, t_1, t_2; \beta, \varepsilon)$ denote the conditional probability analogous to the conditional probability $p^n(y, t_1, t_2; \beta, \varepsilon)$ of Section 3, except that the process $\{\tilde{X}_{\tilde{x}}^n(t), t \in [0, 1]\}$ replaces the process $\{X^n(t), t \in [0, 1]\}$ there. In other words

$$\tilde{p}_{\tilde{x}}^n(y, t_1, t_2; \beta, \varepsilon) \doteq P_{\tilde{x}; y, t_1}^n \left\{ \sup_{s \in [t_1, t_2]} \|\tilde{X}_{\tilde{x}}^n(s) - (s - t_1)\beta\| < \varepsilon \right\},$$

where $P_{\tilde{x}; y, t_1}^n$ denotes probability conditioned on $\tilde{X}_{\tilde{x}}^n(t_1) = y$. We now apply Propositions 3.7 and 3.8 to conclude, for each $\tilde{x} \in \mathcal{S}^\infty$, the existence of a convex function $L(\tilde{x}, \cdot)$ mapping $\Gamma(\tilde{x})$ into $[0, \infty)$ such that for each $\beta \in \Gamma(\tilde{x})$

$$\begin{aligned} & \lim_{\varepsilon \rightarrow 0} \lim_{\delta \rightarrow 0} \liminf_{n \rightarrow \infty} \inf_{\{y \in \mathcal{S}_{\tilde{x}}^n : \|y\| \leq \delta\}} \left(-\frac{1}{n} \log \tilde{p}_{\tilde{x}}^n(y, t_1, t_2; \beta, \varepsilon) \right) \\ (4.6) \quad & = \lim_{\varepsilon \rightarrow 0} \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\{y \in \mathcal{S}_{\tilde{x}}^n : \|y\| \leq \delta\}} \left(-\frac{1}{n} \log \tilde{p}_{\tilde{x}}^n(y, t_1, t_2; \beta, \varepsilon) \right) \\ & = (t_2 - t_1) \cdot L(\tilde{x}, \beta). \end{aligned}$$

In these limits, in order to apply Propositions 3.7 and 3.8 directly, the starting point $\tilde{X}_{\tilde{x}}^n(t_1) = y$ had to be close to the origin. For use in Section 6, it is convenient to restate the limits when the problem is translated so that the starting point $\tilde{X}_{\tilde{x}}^n(t_1) = y$ is away from the origin. We have the following fact.

Proposition 4.2. *We assume Conditions 3.4 and 4.1. Then for each $\tilde{x} \in \mathcal{S}^\infty$ and $z \in \Gamma(\tilde{x})$ there exists a convex function $L(\tilde{x}, \cdot)$ mapping $\Gamma(\tilde{x})$ into $[0, \infty)$ such that for any numbers t_1 and t_2 satisfying $0 \leq t_1 \leq t_2 \leq 1$, $\beta \in \Gamma(\tilde{x})$, and $\varepsilon > 0$*

$$\begin{aligned} & \lim_{\varepsilon \rightarrow 0} \lim_{\delta \rightarrow 0} \liminf_{n \rightarrow \infty} \inf_{\{y \in \mathcal{S}_{\tilde{x}}^n : \|y - z\| \leq \delta\}} \left(-\frac{1}{n} \log P_{\tilde{x}; y, t_1}^n \left\{ \sup_{s \in [t_1, t_2]} \|\tilde{X}_{\tilde{x}}^n(s) - z - (s - t_1)\beta\| < \varepsilon \right\} \right) \\ (4.7) \quad & = \lim_{\varepsilon \rightarrow 0} \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\{y \in \mathcal{S}_{\tilde{x}}^n : \|y - z\| \leq \delta\}} \left(-\frac{1}{n} \log P_{\tilde{x}; y, t_1}^n \left\{ \sup_{s \in [t_1, t_2]} \|\tilde{X}_{\tilde{x}}^n(s) - z - (s - t_1)\beta\| < \varepsilon \right\} \right) \\ & = (t_2 - t_1) \cdot L(\tilde{x}, \beta). \end{aligned}$$

For $\tilde{x} \in \mathcal{S}^\infty$ and $\beta \in \Gamma(\tilde{x})$ the quantity $L(\tilde{x}, \beta)$ is uniquely determined by the limits in formula (4.7).

Proof. If $z = 0$, then there is nothing to prove. On the other hand, if z is a nonzero point in $\Gamma(\tilde{x})$, then there exists a sequence of points $\{z^n, n \in \mathbb{N}\}$

such that each $z^n \in \Gamma(\tilde{x}) \cap \mathcal{S}^n$ and $\lim_{n \rightarrow \infty} z^n = z$. The proposition now follows from the translation invariance in the z variable for $z \in \Gamma(\tilde{x})$, from Propositions 3.7 and 3.8, and from the fact that the limit $n \rightarrow \infty$ is taken before the limits $\delta \rightarrow 0$ and $\varepsilon \rightarrow 0$. \square

Before we can state the large deviation principle, we need a few more definitions. We topologize the function space $\mathcal{D}([0, 1] : \mathbb{R}^d)$ by the Skorohod metric, with respect to which it is a complete separable metric space [2, Sect. 14]. Let $\mathcal{F}([0, 1] : \mathbb{R}^d)$ denote the subset of $\mathcal{D}([0, 1] : \mathbb{R}^d)$ consisting of piecewise linear functions φ mapping $[0, 1]$ into \mathbb{R}^d and having only finitely many discontinuities. For $t \in (0, 1)$ not a point of discontinuity of φ , $\dot{\varphi}(t)$ denotes the derivative of φ with respect to t . Let $\mathcal{F}^\infty([0, 1] : \mathbb{R}^d)$ denote the subset of $\mathcal{F}([0, 1] : \mathbb{R}^d)$ consisting of functions φ for which $\varphi(s) \in \mathcal{S}^\infty$ for every $s \in [0, 1]$. For $x \in \mathcal{S}^\infty$ and $\varphi \in \mathcal{F}^\infty([0, 1] : \mathbb{R}^d)$ satisfying $\varphi(0) = x$, we define the quantity

$$(4.8) \quad I_x(\varphi) \doteq \int_0^1 L(\varphi(t), \dot{\varphi}(t)) dt.$$

Since φ is piecewise linear, for all but finitely many $t \in [0, 1]$ we have $\dot{\varphi}(t) \in \Gamma(\varphi(t))$. Hence $I_x(\varphi)$ is well defined. In addition, since $L(\xi, \beta)$ is finite for $\xi \in \mathcal{S}^\infty$ when $\beta \in \Gamma(\xi)$, $I_x(\varphi)$ is finite for $x \in \mathcal{S}^\infty$ and $\varphi \in \mathcal{F}^\infty([0, 1] : \mathbb{R}^d)$ satisfying $\varphi(0) = x$. For all other $x \in \mathbb{R}^d$ and $\varphi \in \mathcal{D}([0, 1] : \mathbb{R}^d)$ we set $I_x(\varphi) \doteq \infty$. In particular

$$I_x(\varphi) = \infty \text{ for all } \varphi \in \mathcal{F}([0, 1] : \mathbb{R}^d) \setminus \mathcal{F}^\infty([0, 1] : \mathbb{R}^d).$$

Given any set A in $\mathcal{D}([0, 1] : \mathbb{R}^d)$ we write $I_x(A)$ to denote $\inf_{\psi \in A} I_x(\psi)$.

Let $\sigma(\cdot, \cdot)$ denote the Skorohod metric on $\mathcal{D}([0, 1] : \mathbb{R}^d)$, and for $\psi \in \mathcal{D}([0, 1] : \mathbb{R}^d)$ and $\alpha > 0$ we define the open ball

$$B_\sigma(\psi, \alpha) \doteq \{\zeta \in \mathcal{D}([0, 1] : \mathbb{R}^d) : \sigma(\zeta, \psi) < \alpha\}.$$

The rate function in the large deviation principle to be stated just below will be the lower semicontinuous regularization of the function I_x . Specifically, for $x \in \mathbb{R}^d$ and any function $\psi \in \mathcal{D}([0, 1] : \mathbb{R}^d)$ we define the quantity

$$(4.9) \quad J_x(\psi) \doteq \lim_{\alpha \rightarrow 0} \left(\inf_{\{y \in \mathbb{R}^d : \|y-x\| < \alpha\}} I_y(B_\sigma(\psi, \alpha)) \right).$$

Finally, for any set A in $\mathcal{D}([0, 1] : \mathbb{R}^d)$ we write $J_x(A)$ to denote $\inf_{\psi \in A} J_x(\psi)$.

Here is the main theorem of the paper, the large deviation principle for the jump Markov processes $\{X^n(t), t \in [0, 1], n \in \mathbb{N}\}$ having state space \mathcal{S} and jump intensity function $r(x, v)$. While the theorem is formulated for the space $\mathcal{D}([0, 1] : \mathbb{R}^d)$, there is an analogous statement for the space $\mathcal{D}([0, T] : \mathbb{R}^d)$ for any $T \in (0, \infty)$.

Theorem 4.3. *We assume Conditions 3.4 and 4.1. For each $x \in \mathcal{S}^\infty$ let $L(x, \cdot)$ be the function mapping $\Gamma(x)$ into $[0, \infty)$ that satisfies the limits in equation (4.7). For $\psi \in \mathcal{D}([0, 1] : \mathbb{R}^d)$ we define the function $J_x(\psi)$ by formulas (4.8) and (4.9). Then for each $x \in \mathcal{S}^\infty$ the following conclusions hold.*

(a) The function J_x defined in equation (4.9) is nonnegative and lower semi-continuous on $\mathcal{D}([0, 1]: \mathbb{R}^d)$. Furthermore for any compact set V in \mathbb{R}^d and any $M \in [0, \infty)$ the set

$$\bigcup_{x \in V} \{ \psi \in \mathcal{D}([0, 1]: \mathbb{R}^d) : J_x(\psi) \leq M \}$$

is a compact set in $\mathcal{D}([0, 1]: \mathbb{R}^d)$.

(b) For $\varphi \in \mathcal{I}^\infty([0, 1]: \mathbb{R}^d)$ and $x \doteq \varphi(0) \in \mathbb{R}^d$, we have $J_x(\varphi) = I_x(\varphi) < \infty$. For $\varphi \in \mathcal{I}([0, 1]: \mathbb{R}^d) \setminus \mathcal{I}^\infty([0, 1]: \mathbb{R}^d)$ and $x \doteq \varphi(0) \in \mathbb{R}^d$, we have $J_x(\varphi) = I_x(\varphi) = \infty$.

(c) For any open set G in $\mathcal{D}([0, 1]: \mathbb{R}^d)$ and each $x \in \mathbb{R}^d$ we have the large deviation lower bound

$$\lim_{\delta \rightarrow 0} \liminf_{n \rightarrow \infty} \inf_{\{y \in \mathcal{I}^n : \|y-x\| \leq \delta\}} \frac{1}{n} \log P_{y,0}^n \{X^n \in G\} \geq -J_x(G).$$

(d) For any closed set F in $\mathcal{D}([0, 1]: \mathbb{R}^d)$ and each $x \in \mathbb{R}^d$ we have the large deviation upper bound

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\{y \in \mathcal{I}^n : \|y-x\| \leq \delta\}} \frac{1}{n} \log P_{y,0}^n \{X^n \in F\} \leq -J_x(F).$$

We conjecture that if ψ is absolutely continuous, ψ satisfies $\psi(0) = x$, and $\psi(s)$ lies in \mathcal{I}^∞ for each $s \in [0, 1]$, then $J_x(\psi)$ is given by a formula as in equation (4.8) and that for all other $\psi \in \mathcal{D}([0, 1]: \mathbb{R}^d)$ $J_x(\psi) = \infty$. This conjecture will be considered in the sequel to this paper.

We end this section by describing how one can apply this large deviation principle to a queueing system with finite buffers.

Example 4.4. The purpose of the present example is to indicate the variety of other queueing systems for which the large deviation principle can be proved by using Theorem 4.3. Let M and N be two positive integers. For $n \in \mathbb{N}$ we consider a queueing system consisting of two tandem queues, the first of which has a finite buffer size Mn and the second of which has a finite buffer size Nn . We will consider the asymptotic behavior of the system as $n \rightarrow \infty$. If queue 1 is not filled, then new customers enter queue 1 at the rate $\alpha > 0$ and are served by queue 1 at the rate $\beta > 0$. After being served in queue 1, a customer enters queue 2 as long as queue 2 is not filled. Customers are served by queue 2 at the rate $\gamma > 0$. If queue 1 is filled at the time of arrival of a new customer, then the customer is rejected and leaves the system. If after being served by queue 1 a customer finds that queue 2 is filled, then the customer recycles and enters queue 1 again.

We model the system by a jump Markov process $\{X(t), t \in [0, \infty)\}$, where $X(t) = (X_1(t), X_2(t))$ and for each $i \in \{1, 2\}$ $X_i(t)$ denotes the sum of the number of customers waiting in queue i and the number of customers being served in queue i . The state space of $X(t)$ is the n -dependent rectangle

$$\mathcal{I}^n \doteq \{(j, k) \in \mathbb{Z}^2 : j \in [0, Mn], k \in [0, Nn]\}.$$

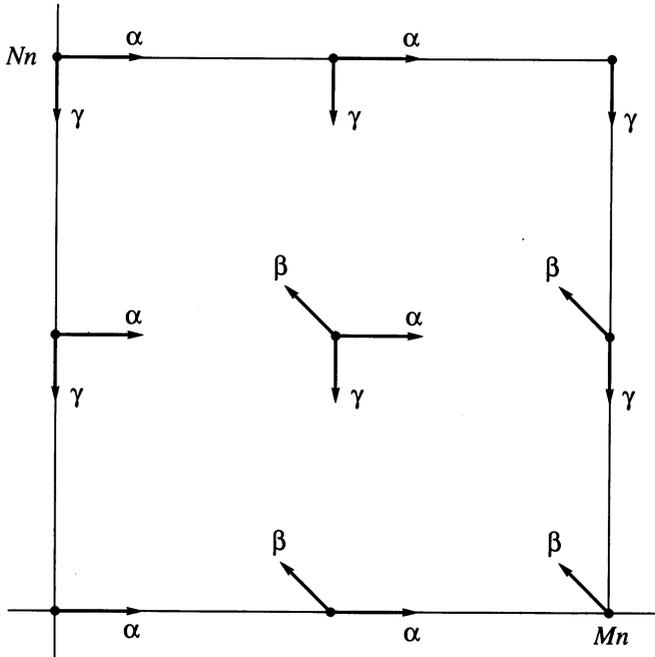


FIGURE 4. Jump intensities for the queueing system with finite buffers.

The possible jumps and corresponding jump intensities are indicated in Figure 4.

For $n \in \mathbb{N}$ we consider the scaled jump Markov process $\{X^n(t), t \in [0, 1]\}$ defined by the formula

$$X^n(t) \doteq \frac{1}{n}X(nt)$$

and having the state space

$$(\mathcal{S}^n)^n \doteq \left\{ \frac{1}{n}y : y \in \mathcal{S}^n \right\}.$$

For this process the sizes of the buffers scale to M and N . Although the queueing system with buffers does not satisfy either Condition 3.1 or Condition 4.1, the large deviation principle stated in Theorem 4.3 can be applied, in combination with localization arguments, to prove the analogous large deviation principle for the sequence of scaled jump Markov processes $\{X^n(t), t \in [0, 1], n \in \mathbb{N}\}$.

Consider the limit state space

$$\begin{aligned} \mathcal{S}^\infty &\doteq \{y \in \mathbb{R}^2 : y = \lim_{n \rightarrow \infty} y^n \text{ for some sequence } \{y^n, n \in \mathbb{N}\} \\ &\qquad\qquad\qquad \text{with each } y^n \in \mathcal{S}^n\} \\ &= [0, M] \times [0, N], \end{aligned}$$

which is a rectangle containing the four vertices $(0, 0)$, $(M, 0)$, $(0, N)$, and (M, N) . The localization arguments needed to prove the large deviation principle are standard and seem to have originated in [1]. We take open neighborhoods of each of the four vertices of the limit state space which do not contain any other vertex and whose union is the whole rectangle $[0, M] \times [0, N]$. We then start the process near a given vertex, say $(0, 0)$, and run it either until it leaves the neighborhood of $(0, 0)$ or until time 1, whichever is smaller. The large deviation principle for the stopped process holds by Theorem 4.3. We now keep repeating the procedure. Each time that the process is stopped before time 1, when stopped it lies in a different neighborhood of another vertex. We run the process again either until it leaves this new neighborhood or until time 1, whichever is smaller. The overall large deviation principle is obtained by patching these local large deviation principles together. This completes the discussion of the tandem queueing system with finite buffers. \square

The proof of Theorem 4.3 is given in the next section. It depends on a local large deviation principle that is proved in Section 6.

5. PROOF OF THE LARGE DEVIATION PRINCIPLE IN THEOREM 4.3

We recall the sets of functions $\mathcal{S}([0, 1] : \mathbb{R}^d)$ and $\mathcal{S}^\infty([0, 1] : \mathbb{R}^d)$, which were defined before Theorem 4.3. In order to prove the theorem, we need the following proposition, the proof of which is deferred to the next section of this paper. The proposition states a local large deviation principle for balls centered at $\varphi \in \mathcal{S}([0, 1] : \mathbb{R}^d)$.

Proposition 5.1. *We assume Conditions 3.4 and 4.1. Let φ be any function in $\mathcal{S}([0, 1] : \mathbb{R}^d)$ and set $x \doteq \varphi(0)$. Then*

$$\begin{aligned} &\lim_{\varepsilon \rightarrow 0} \lim_{\delta \rightarrow 0} \liminf_{n \rightarrow \infty} \inf_{\{y \in \mathcal{S}^n : \|y-x\| \leq \delta\}} \frac{1}{n} \log P_{y,0}^n \left\{ \sup_{t \in [0,1]} \|X^n(t) - \varphi(t)\| < \varepsilon \right\} \\ &= \lim_{\varepsilon \rightarrow 0} \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\{y \in \mathcal{S}^n : \|y-x\| \leq \delta\}} \frac{1}{n} \log P_{y,0}^n \left\{ \sup_{t \in [0,1]} \|X^n(t) - \varphi(t)\| < \varepsilon \right\} \\ &= -I_x(\varphi). \end{aligned}$$

Before proceeding with the proof of Theorem 4.3, we recall some definitions and facts. The space $\mathcal{D}([0, 1] : \mathbb{R}^d)$ is a complete separable metric space with respect to the Skorohod metric $\sigma(\cdot, \cdot)$. We denote the supremum norm metric

by $\rho(\cdot, \cdot)$, and for $\psi \in \mathcal{D}([0, 1] : \mathbb{R}^d)$ and $\varepsilon > 0$ we define the open balls

$$B_\sigma(\psi, \varepsilon) \doteq \{ \zeta \in \mathcal{D}([0, 1] : \mathbb{R}^d) : \sigma(\zeta, \psi) < \varepsilon \}$$

and

$$B_\rho(\psi, \varepsilon) \doteq \{ \zeta \in \mathcal{D}([0, 1] : \mathbb{R}^d) : \rho(\zeta, \psi) < \varepsilon \}.$$

The symbols $\bar{B}_\rho(\psi, \varepsilon)$ and $\bar{B}_\sigma(\psi, \varepsilon)$ denote the corresponding closed balls. We have the following properties: for any $\psi \in \mathcal{D}([0, 1] : \mathbb{R}^d)$ and $\varepsilon > 0$

$$B_\rho(\psi, \varepsilon) \subset B_\sigma(\psi, \varepsilon)$$

and for any $\psi \in \mathcal{E}([0, 1] : \mathbb{R}^d)$, each $\varepsilon > 0$, and all sufficiently small $\alpha > 0$ depending on ε

$$B_\sigma(\psi, \alpha) \subset B_\rho(\psi, \varepsilon).$$

The proofs of parts (a) and (d) of Theorem 4.3 rely on some results in [7]. The main theorem in that paper, Theorem 1.1, proves a large deviation upper bound for a class of locally infinitely divisible Markov processes taking values in $\mathcal{D}([0, 1] : \mathbb{R}^d)$. This class includes as a special case the jump Markov processes $\{X^n(t), t \in [0, 1], n \in \mathbb{N}\}$ now under consideration. However, the optimal large deviation upper bound stated in Theorem 4.3 in the present paper is better in general than the large deviation upper bound proved in Theorem 1.1 of [7] (cf. Lemma 5.3). The latter theorem is stated in terms of a function $K_x(\psi)$ that is defined for each $\psi \in \mathcal{D}([0, 1] : \mathbb{R}^d)$ and each $x \in \mathbb{R}^d$ and that takes values in $[0, \infty]$. If ψ is absolutely continuous and $\psi(0) = x$, then $K_x(\psi)$ has the form

$$(5.1) \quad K_x(\psi) = \int_0^1 F(\psi(s), \dot{\psi}(s)) ds,$$

where for each $\xi \in \mathbb{R}^d$ and $\beta \in \mathbb{R}^d$ $F(\xi, \beta)$ is defined as the Legendre-Fenchel transform, with respect to the variable β , of an explicitly given function $H(\xi, \alpha)$ of $\xi \in \mathbb{R}^d$ and $\alpha \in \mathbb{R}^d$. In the case of queueing systems, an explicit formula for $F(\xi, \beta)$ is given in Theorem 3.1 of [7]. If ψ is not absolutely continuous or if $\psi(0) \neq x$, then $K_x(\psi) \doteq \infty$.

When specialized to the jump Markov processes $\{X^n(t), t \in [0, 1], n \in \mathbb{N}\}$, part (a) of Theorem 1.1 in [7] gives part (a) of the following theorem. For parts (b) and (c) we introduce the piecewise linear interpolations $\{Y^n(t), t \in [0, 1], n \in \mathbb{N}\}$, defined by the formula

$$(5.2) \quad Y^n(t) \doteq n \left[\left(\frac{j+1}{n} - t \right) \cdot X^n(j/n) + \left(t - \frac{j}{n} \right) \cdot X^n((j+1)/n) \right]$$

for $t \in [j/n, (j+1)/n]$, $j \in \{0, 1, \dots, n-1\}$. These processes take values in $\mathcal{E}([0, 1] : \mathbb{R}^d)$. Part (b), proved in Lemma 2.4 of [7], proves that the jump Markov processes $\{X^n(t), t \in [0, 1], n \in \mathbb{N}\}$ and the piecewise linear interpolations $\{Y^n(t), t \in [0, 1], n \in \mathbb{N}\}$ are superexponentially close. Part (c), proved in Lemma 2.5 of [7], proves that the processes $\{Y^n(t), t \in [0, 1], n \in \mathbb{N}\}$ are exponentially tight.

Theorem 5.2. *Let V be a compact set in \mathbb{R}^d . For any $x \in \mathbb{R}^d$ and $M \in [0, \infty)$ we consider the level set*

$$A(K_x, M) \doteq \{\psi \in \mathcal{D}([0, 1] : \mathbb{R}^d) : K_x(\psi) \leq M\}.$$

The following conclusions hold.

(a) *The set $\bigcup_{x \in V} A(K_x, M)$ is a compact subset of $\mathcal{D}([0, 1] : \mathbb{R}^d)$, and therefore the function mapping*

$$(x, \psi) \in \mathbb{R}^d \times \mathcal{D}([0, 1] : \mathbb{R}^d) \longmapsto K_x(\psi) \in [0, \infty]$$

is lower semicontinuous.

(b) *For any $\beta > 0$*

$$\limsup_{n \rightarrow \infty} \sup_{y \in \mathbb{R}^d} \frac{1}{n} \log P_{y,0}^n \left\{ \sup_{t \in [0,1]} \|X^n(t) - Y^n(t)\| \geq \beta \right\} = -\infty.$$

(c) *Let $B \in (0, \infty)$ be given. Then there exists a compact set Π in $\mathcal{E}([0, 1] : \mathbb{R}^d)$ such that*

$$\limsup_{n \rightarrow \infty} \sup_{y \in V} \frac{1}{n} \log P_{y,0}^n \{Y^n \notin \Pi\} \leq -B.$$

At this point let us recall the functions I_x and J_x , which were defined just before the statement of Theorem 4.3. Parts (b) and (c) of the next lemma show that I_x and J_x are pointwise no less than K_x . This is due to the fact that K_x is an upper rate function but not necessarily the true rate function. The limit in part (a) is needed in the proof of part (b). For $x \in \mathbb{R}^d$ and A a subset of $\mathcal{D}([0, 1] : \mathbb{R}^d)$ $K_x(A)$ denotes $\inf_{\psi \in A} K_x(\psi)$.

Lemma 5.3.

(a) *For any $x \in \mathbb{R}^d$ and closed set F in $\mathcal{D}([0, 1] : \mathbb{R}^d)$*

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \log P_{x,0}^n \{X^n \in F\} \leq -K_x(F).$$

(b) *For all $\varphi \in \mathcal{D}([0, 1] : \mathbb{R}^d)$ and each $x \in \mathbb{R}^d$ $I_x(\varphi) \geq K_x(\varphi)$.*

(c) *For all $\psi \in \mathcal{D}([0, 1] : \mathbb{R}^d)$ and each $x \in \mathbb{R}^d$ $J_x(\psi) \geq K_x(\psi)$.*

(a) This is proved in part (b) of Theorem 1.1 in [7].

(b) Since for any $\varepsilon > 0$ $B_\rho(\varphi, \varepsilon) \subset \overline{B}_\sigma(\varphi, \varepsilon)$, we have from Proposition 5.1 and part (a) of the present lemma

$$\begin{aligned} -I_x(\varphi) &= \lim_{\varepsilon \rightarrow 0} \lim_{\delta \rightarrow 0} \liminf_{n \rightarrow \infty} \inf_{\{y \in \mathcal{S}^n : \|y-x\| \leq \delta\}} \frac{1}{n} \log P_{y,0}^n \left\{ \sup_{t \in [0,1]} \|X^n(t) - \varphi(t)\| < \varepsilon \right\} \\ &\leq \lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \log P_{x,0}^n \{X^n \in \overline{B}_\sigma(\varphi, \varepsilon)\} \\ &\leq -\lim_{\varepsilon \rightarrow 0} K_x(\overline{B}_\sigma(\varphi, \varepsilon)) \\ &= -K_x(\varphi). \end{aligned}$$

The last line is a consequence of the lower semicontinuity of K_x (part (a) of Theorem 5.2). We have proved that $K_x(\varphi) \leq I_x(\varphi)$.

(c) Since $I_x \geq K_x$ on $\mathcal{D}([0, 1] : \mathbb{R}^d)$ and since the function mapping $(x, \psi) \mapsto K_x(\psi)$ is lower semicontinuous (part (a) of Theorem 5.2), we have

$$\begin{aligned} J_x(\psi) &\doteq \lim_{\alpha \rightarrow 0} \left(\inf_{\{y \in \mathbb{R}^d : \|y-x\| < \alpha\}} I_y(B_\sigma(\psi, \alpha)) \right) \\ &\geq \lim_{\alpha \rightarrow 0} \left(\inf_{\{y \in \mathbb{R}^d : \|y-x\| < \alpha\}} K_y(B_\sigma(\psi, \alpha)) \right) \\ &= K_x(\psi). \end{aligned}$$

This proves part (c).

We are now ready to prove part (a) of Theorem 4.3. This states a number of properties of the function J_x .

Proof of part (a) of Theorem 4.3. The nonnegativity of J_x and the lower semicontinuity of the function mapping $(x, \psi) \in \mathbb{R}^d \times \mathcal{D}([0, 1] : \mathbb{R}^d) \mapsto J_x(\psi)$ both follow from the definition of J_x in terms of the nonnegative function I_x . For any $x \in \mathbb{R}^d$ and $M \in [0, \infty)$ we consider the level set

$$A(J_x, M) \doteq \{\psi \in \mathcal{D}([0, 1] : \mathbb{R}^d) : J_x(\psi) \leq M\}$$

and prove that for any compact set V in \mathbb{R}^d the set $\bigcup_{x \in V} A(J_x, M)$ is a compact set in $\mathcal{D}([0, 1] : \mathbb{R}^d)$. The set $\bigcup_{x \in V} A(J_x, M)$ is closed because of the lower semicontinuity of the function mapping $(x, \psi) \mapsto J_x(\psi)$. Since for each $\psi \in \mathcal{D}([0, 1] : \mathbb{R}^d)$ and $x \in \mathbb{R}^d$ $J_x(\psi) \geq K_x(\psi)$ (part (c) of Lemma 5.3), it follows that

$$\bigcup_{x \in V} A(J_x, M) \subset \bigcup_{x \in V} A(K_x, M).$$

Since $\bigcup_{x \in V} A(K_x, M)$ is a compact subset of $\mathcal{D}([0, 1] : \mathbb{R}^d)$ (part (a) of Theorem 5.2), we conclude that $\bigcup_{x \in V} A(J_x, M)$ is also a compact subset of $\mathcal{D}([0, 1] : \mathbb{R}^d)$. \square

We next prove the large deviation lower bound stated in part (c) of Theorem 4.3. Parts (b) and (d) will be proved afterwards.

Proof of part (c) of Theorem 4.3. If $J_x(G) = \infty$, then there is nothing to prove. So we assume that $J_x(G) < \infty$. Let $\gamma > 0$ be given. Then there exists $\psi \in G$ such that

$$J_x(G) + \gamma \geq J_x(\psi).$$

By the definition of J_x as the lower semicontinuous regularization of I_x , for any $\alpha > 0$ there exists a point $y_\alpha \in \mathbb{R}^d$ satisfying $\|y_\alpha - x\| \leq \alpha$ and a function $\varphi_\alpha \in B_\sigma(\psi, \alpha)$ such that

$$\infty > J_x(\psi) + \gamma \geq I_{y_\alpha}(B_\sigma(\psi, \alpha)) \geq I_{y_\alpha}(\varphi_\alpha) - \gamma.$$

Since G is open there exists $\alpha_1 > 0$ such that $\varphi_\alpha \in G$ for all $\alpha \in (0, \alpha_1)$, and for all $\alpha \in (0, \alpha_1)$ there exists $\lambda_\alpha > 0$ such that for all $\lambda \in (0, \lambda_\alpha)$

$$B_\rho(\varphi_\alpha, \lambda) \subset B_\sigma(\varphi_\alpha, \lambda) \subset G.$$

Given $\Delta > 0$ there exists $\alpha_2 \in (0, \alpha_1)$ and $\delta_1 > 0$ such that for all $\alpha \in (0, \alpha_2)$ and all $\delta \in (0, \delta_1)$

$$\{y \in \mathcal{S}^n : \|y - x\| \leq \delta\} \subset \{y \in \mathcal{S}^n : \|y - y_\alpha\| \leq \Delta\}.$$

Hence by Proposition 5.1 we have for all $\alpha \in (0, \alpha_2)$

$$\begin{aligned} \lim_{\delta \rightarrow 0} \liminf_{n \rightarrow \infty} \inf_{\{y \in \mathcal{S}^n : \|y-x\| \leq \delta\}} \frac{1}{n} \log P_{y,0}^n \{X^n \in G\} \\ \geq \lim_{\lambda \rightarrow 0} \lim_{\delta \rightarrow 0} \liminf_{n \rightarrow \infty} \inf_{\{y \in \mathcal{S}^n : \|y-x\| \leq \delta\}} \frac{1}{n} \log P_{y,0}^n \{X^n \in B_\rho(\varphi_\alpha, \lambda)\} \\ \geq \lim_{\lambda \rightarrow 0} \lim_{\Delta \rightarrow 0} \liminf_{n \rightarrow \infty} \inf_{\{y \in \mathcal{S}^n : \|y-y_\alpha\| \leq \Delta\}} \frac{1}{n} \log P_{y,0}^n \{X^n \in B_\rho(\varphi_\alpha, \lambda)\} \\ = -I_{y_\alpha}(\varphi_\alpha) \\ \geq -J_x(G) - 3\gamma. \end{aligned}$$

Taking $\gamma \rightarrow 0$, we obtain the desired lower bound. This completes the proof of part (c) of Theorem 4.3. \square

We now prove part (b) of Theorem 4.3.

Proof of part (b) of Theorem 4.3. The definition of J_x shows that $J_x(\varphi) \leq I_x(\varphi)$ for all $\varphi \in \mathcal{F}([0, 1] : \mathbb{R}^d)$ and each $x \in \mathbb{R}^d$. We now prove the reverse inequality. Since $\varphi \in \mathcal{F}([0, 1] : \mathbb{R}^d)$ is continuous, for each $\varepsilon > 0$ and all sufficiently small $\alpha > 0$ depending on ε we have $B_\sigma(\varphi, \alpha) \subset B_\rho(\varphi, \varepsilon)$. Proposition 5.1, the large deviation lower bound just proved, and the lower semicontinuity of J_x on $\mathcal{D}([0, 1] : \mathbb{R}^d)$ show that

$$\begin{aligned} -I_x(\varphi) &= \lim_{\varepsilon \rightarrow 0} \lim_{\delta \rightarrow 0} \liminf_{n \rightarrow \infty} \inf_{\{y \in \mathcal{S}^n : \|y-x\| \leq \delta\}} \frac{1}{n} \log P_{y,0}^n \left\{ \sup_{t \in [0,1]} \|X^n(t) - \varphi(t)\| < \varepsilon \right\} \\ &\geq \lim_{\alpha \rightarrow 0} \lim_{\delta \rightarrow 0} \liminf_{n \rightarrow \infty} \inf_{\{y \in \mathcal{S}^n : \|y-x\| \leq \delta\}} \frac{1}{n} \log P_{y,0}^n \{X^n \in B_\sigma(\varphi, \alpha)\} \\ &\geq -\limsup_{\alpha \rightarrow 0} J_x(B_\sigma(\varphi, \alpha)) \\ &= -J_x(\varphi). \end{aligned}$$

We have proved that $J_x(\varphi) \geq I_x(\varphi)$. It follows that $J_x(\varphi) = I_x(\varphi)$ for all $\varphi \in \mathcal{F}([0, 1] : \mathbb{R}^d)$. As we noted after the definition of $I_x(\varphi)$ in equation (4.8), $I_x(\varphi) < \infty$ if $\varphi \in \mathcal{F}^\infty([0, 1] : \mathbb{R}^d)$ and $I_x(\varphi) = \infty$ if $\varphi \in \mathcal{F}([0, 1] : \mathbb{R}^d) \setminus \mathcal{F}^\infty([0, 1] : \mathbb{R}^d)$. \square

We are now ready to prove the large deviation upper bound stated in part (d) of Theorem 4.3.

Proof of part (d) of Theorem 4.3. Let us assume that we can prove that for any closed set F in $\mathcal{D}([0, 1]: \mathbb{R}^d)$ and each $x \in \mathbb{R}^d$

$$(5.3) \quad \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\{y \in \mathcal{S}^n : \|y-x\| \leq \delta\}} \frac{1}{n} \log P_{y,0}^n \{Y^n \in F\} \leq -J_x(F),$$

where $\{Y^n(t), t \in [0, 1], n \in \mathbb{N}\}$ are the piecewise linear interpolations of the jump Markov processes $\{X^n(t), t \in [0, 1], n \in \mathbb{N}\}$ defined in equation (5.2). We will prove the limit (5.3) later. For any $y \in \mathbb{R}^d$ and any $\beta > 0$

$$\begin{aligned} P_{y,0}^n \{X^n \in F\} &\leq P_{y,0}^n \{X^n \in F, \rho(Y^n, X^n) \leq \beta\} + P_{y,0}^n \{\rho(X^n, Y^n) \geq \beta\} \\ &\leq P_{y,0}^n \{\sigma(Y^n, F) \leq \beta\} + P_{y,0}^n \{\rho(X^n, Y^n) \geq \beta\}. \end{aligned}$$

Hence by the limit (5.3) and the superexponential estimate in part (b) of Theorem 5.2

$$(5.4) \quad \begin{aligned} &\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\{y \in \mathcal{S}^n : \|y-x\| \leq \delta\}} \frac{1}{n} \log P_{y,0}^n \{X^n \in F\} \\ &\leq \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\{y \in \mathcal{S}^n : \|y-x\| \leq \delta\}} \frac{1}{n} \log P_{y,0}^n \{\sigma(Y^n, F) \leq \beta\} \\ &\leq -J_x \left(\{\psi \in \mathcal{D}([0, 1]: \mathbb{R}^d) : \sigma(\psi, F) \leq \beta\} \right). \end{aligned}$$

Since J_x has compact level sets (part (a) of Theorem 4.3), we have

$$\lim_{\beta \rightarrow 0} J_x \left(\{\psi \in \mathcal{D}([0, 1]: \mathbb{R}^d) : \sigma(\psi, F) \leq \beta\} \right) = J_x(F).$$

Thus letting $\beta \rightarrow 0$ in formula (5.4), we obtain part (d) of Theorem 4.3.

We now prove the limit (5.3) for F a closed set in $\mathcal{D}([0, 1]: \mathbb{R}^d)$. For such a set, $F \cap \mathcal{E}([0, 1]: \mathbb{R}^d)$ is a closed set in $\mathcal{E}([0, 1]: \mathbb{R}^d)$. Since the processes $\{Y^n(t), t \in [0, 1], n \in \mathbb{N}\}$ take values in $\mathcal{E}([0, 1]: \mathbb{R}^d)$ and satisfy the exponential tightness property expressed in part (c) of Theorem 5.2, it suffices to prove the limit (5.3) for F a compact subset of $\mathcal{E}([0, 1]: \mathbb{R}^d)$ [3, Lemma 2.1.5]. We give the proof assuming that $J_x(F) < \infty$. The routine modifications needed in the case $J_x(F) = \infty$ are omitted.

The superexponential estimate in part (b) of Theorem 5.2 allows us to carry over to the processes $\{Y^n(t), t \in [0, 1], n \in \mathbb{N}\}$ the local large deviation principle stated in Proposition 5.1 for the processes $\{X^n(t), t \in [0, 1], n \in \mathbb{N}\}$. Thus let φ be any function in $\mathcal{T}([0, 1]: \mathbb{R}^d)$ and set $x \doteq \varphi(0)$. Then

$$\begin{aligned} &\lim_{\varepsilon \rightarrow 0} \lim_{\delta \rightarrow 0} \liminf_{n \rightarrow \infty} \inf_{\{y \in \mathcal{S}^n : \|y-x\| \leq \delta\}} \frac{1}{n} \log P_{y,0}^n \left\{ \sup_{t \in [0, 1]} \|Y^n(t) - \varphi(t)\| < \varepsilon \right\} \\ &= \lim_{\varepsilon \rightarrow 0} \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\{y \in \mathcal{S}^n : \|y-x\| \leq \delta\}} \frac{1}{n} \log P_{y,0}^n \left\{ \sup_{t \in [0, 1]} \|Y^n(t) - \varphi(t)\| < \varepsilon \right\} \\ &= -I_x(\varphi). \end{aligned}$$

The function $I_x(\varphi)$ is finite for $\varphi \in \mathcal{T}^\infty([0, 1]: \mathbb{R}^d)$ and is infinite for $\varphi \in \mathcal{T}([0, 1]: \mathbb{R}^d) \setminus \mathcal{T}^\infty([0, 1]: \mathbb{R}^d)$, and for any $\varphi \in \mathcal{T}([0, 1]: \mathbb{R}^d)$ $I_x(\varphi) = J_x(\varphi)$ (part (b) of Theorem 4.3). Hence given $\gamma > 0$, for any $\varphi \in$

$\mathcal{F}^\infty([0, 1]: \mathbb{R}^d)$ there exists $\varepsilon_\varphi > 0$ such that for all $\varepsilon \in (0, \varepsilon_\varphi)$

$$\begin{aligned} \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\{y \in \mathcal{S}^n : \|y-x\| \leq \delta\}} \frac{1}{n} \log P_{y,0}^n \left\{ \sup_{t \in [0,1]} \|Y^n(t) - \varphi(t)\| < \varepsilon \right\} \\ \leq -I_x(\varphi) + \gamma = -J_x(\varphi) + \gamma < \infty, \end{aligned}$$

and for any $\varphi \in \mathcal{F}([0, 1]: \mathbb{R}^d) \setminus \mathcal{F}^\infty([0, 1]: \mathbb{R}^d)$ there exists $\varepsilon_\varphi > 0$ such that for all $\varepsilon \in (0, \varepsilon_\varphi)$

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\{y \in \mathcal{S}^n : \|y-x\| \leq \delta\}} \frac{1}{n} \log P_{y,0}^n \left\{ \sup_{t \in [0,1]} \|Y^n(t) - \varphi(t)\| < \varepsilon \right\} \leq -J_x(F).$$

Since $\mathcal{F}([0, 1]: \mathbb{R}^d)$ is a dense subset of $\mathcal{E}([0, 1]: \mathbb{R}^d)$ we can write

$$F \subset \bigcup_{\varphi \in F \cap \mathcal{F}([0,1]:\mathbb{R}^d)} B_\rho(\varphi, \varepsilon_\varphi).$$

The compactness of F guarantees that there exist $r \in \mathbb{N}$ and elements $\{\varphi_i, i = 1, 2, \dots, r\}$ in $F \cap \mathcal{F}([0, 1]: \mathbb{R}^d)$ such that

$$F \subset \bigcup_{i=1}^r B_\rho(\varphi_i, \varepsilon_i), \text{ where } \varepsilon_i \doteq \varepsilon_{\varphi_i}.$$

Thus

$$\begin{aligned} \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\{y \in \mathcal{S}^n : \|y-x\| \leq \delta\}} \frac{1}{n} \log P_{y,0}^n \{Y^n \in F\} \\ \leq \max_{i=1,2,\dots,r} \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\{y \in \mathcal{S}^n : \|y-x\| \leq \delta\}} \frac{1}{n} \log P_{y,0}^n \{Y^n \in B_\rho(\varphi_i, \varepsilon_i)\} \\ \leq - \left(\min_{i=1,2,\dots,r} (J_x(\varphi_i) - \gamma) \right) \wedge J_x(F) \\ \leq -J_x(F) + \gamma. \end{aligned}$$

Taking $\gamma \rightarrow 0$, we obtain the limit (5.3). The proof of part (d) of Theorem 4.3 is complete. \square

Except for the proof of the local large deviation principle stated in Proposition 5.1, we have finished the proof of the large deviation principle stated in Theorem 4.3. The local large deviation principle is proved in the next section.

6. PROOF OF THE LOCAL LARGE DEVIATION PRINCIPLE IN PROPOSITION 5.1

All the notational conventions of Sections 3 and 4 will be in effect for this section as well. Thus the queueing model is defined in terms of a fixed set of nonzero vectors $\{w^{(i)}, i = 1, 2, \dots, \kappa\}$ as described in Condition 4.1. This condition also specifies the form of the jump intensity function $r(\xi, v)$. In terms of these vectors we define the facets $\{\mathcal{F}^{(o_1, \dots, o_\kappa)}\}$, where each o_i represents one of the symbols $>, =, <$. For each point \tilde{x} in the limit state space \mathcal{S}^∞ we define in formula (4.4) the facets $\{\tilde{\mathcal{F}}^{(o_1, \dots, o_\kappa(\tilde{x}))}\}$ determined by \tilde{x} .

There is a one-to-one correspondence between the set of all facets $\tilde{\mathcal{F}}^{(\circ_1, \dots, \circ_{\kappa(\tilde{x})})}$ determined by \tilde{x} and the subset of original facets $\mathcal{F}^{(\circ_1, \dots, \circ_{\kappa})}$ that either contain \tilde{x} or are adjacent to \tilde{x} .

For ξ lying in a facet $\tilde{\mathcal{F}}^{(\circ_1, \dots, \circ_{\kappa(\tilde{x})})}$ and for $v \in \mathbb{Z}^d$ the local jump intensity function $\tilde{r}(\tilde{x}; \xi, v)$ has been defined by the formula

$$(6.1) \quad \tilde{r}(\tilde{x}; \xi, v) \doteq \lim r(z, v).$$

The limit is taken as $z \rightarrow \tilde{x}$ with $z - \tilde{x}$ lying in the same facet $\tilde{\mathcal{F}}^{(\circ_1, \dots, \circ_{\kappa(\tilde{x})})}$ as ξ . We need the following bound, which is a consequence of formula (4.1). For any $\tilde{x} \in \mathcal{S}^\infty$, $\xi \in \tilde{\mathcal{F}}^{(\circ_1, \dots, \circ_{\kappa(\tilde{x})})}$, and $v \in \mathbb{Z}^d$ satisfying $\tilde{r}(\tilde{x}; \xi, v) > 0$ there exists a constant $C \in [1, \infty)$ with the property that

$$(6.2) \quad 1/C \leq \tilde{r}(\tilde{x}; \xi, v) \leq C.$$

For $n \in \mathbb{N}$ and $\tilde{x} \in \mathcal{S}^\infty$ we recall the stationary jump Markov process $\{\tilde{X}_{\tilde{x}}^n(t), t \in [0, 1]\}$ having the state space $\mathcal{S}_{\tilde{x}}^n$ and the local jump intensity function $\tilde{r}(\tilde{x}; \cdot, \cdot)$. In terms of this process we have for β in the set $\Gamma(\tilde{x})$ the limits in formula (4.7). These limits uniquely define the function $L(\tilde{x}, \beta)$. A uniform continuity property of this function is given in the next lemma, which is needed in the proof of Proposition 5.1. The lemma follows directly from Lemma 3.10 applied to the family $r_\lambda(\cdot, \cdot) \doteq \tilde{r}(\tilde{x}; \cdot, \cdot)$ and from the bound (6.2).

Lemma 6.1. *We assume Conditions 3.4 and 4.1. For each $\tilde{x} \in \mathcal{S}^\infty$ and $\beta \in \Gamma(\tilde{x})$ let $L(\tilde{x}, \beta)$ be the function defined by the limits in formula (4.7). Then for each sequence of symbols $(\circ_1, \dots, \circ_{\kappa})$ and any point $y \in \mathcal{F}^{(\circ_1, \dots, \circ_{\kappa})}$ the restriction of $L(\cdot, \cdot)$ to the set $\mathcal{F}^{(\circ_1, \dots, \circ_{\kappa})} \times \Gamma(y)$ is uniformly continuous on bounded subsets of $\mathcal{F}^{(\circ_1, \dots, \circ_{\kappa})} \times \Gamma(y)$.*

We are now ready to prove the local large deviation principle stated in Proposition 5.1.

Proof of Proposition 5.1. We first suppose that φ is an element of $\mathcal{F}([0, 1] : \mathbb{R}^d) \setminus \mathcal{S}^\infty([0, 1] : \mathbb{R}^d)$. Then there exists $s \in [0, 1]$ so that the distance from $\varphi(s)$ to the limit state space \mathcal{S}^∞ is positive. In this case, for all sufficiently small $\varepsilon > 0$, all sufficiently large $n \in \mathbb{N}$, and any $y \in \mathcal{S}^n$ the probability

$$P_{y,0}^n \left\{ \sup_{t \in [0,1]} \|X^n(t) - \varphi(t)\| < \varepsilon \right\}$$

equals 0. Since $I_x(\varphi) = \infty$, the limit in Proposition 5.1 is valid.

We now prove the limit in Proposition 5.1 for a function $\varphi \in \mathcal{S}^\infty([0, 1] : \mathbb{R}^d)$. There exists a positive integer N and a sequence $0 = t^{(0)} < t^{(1)} < t^{(2)} < \dots < t^{(N)} = 1$ representing the times at which the derivative of φ is discontinuous. We define the compact sets

$$\tilde{V} \doteq \{\varphi(t) : t \in [0, 1]\} \cup \{\dot{\varphi}(t) : t \in [0, 1] \setminus \{t^{(0)}, t^{(1)}, \dots, t^{(N)}\}\}$$

and

$$V \doteq \left\{ y \in \mathbb{R}^d : \|x - y\| \leq 1 \text{ for some } x \in \tilde{V} \right\}.$$

Now let $\gamma \in (0, 1)$ be given. According to part (a) of Lemma 6.1, we can choose $\Delta_1 > 0$ with the following property. If $\Delta \in (0, \Delta_1)$, if both z and ξ lie in the set $V \cap \mathcal{F}^{(\circ_1, \dots, \circ_\kappa)}$ for some choice of $\circ_1, \dots, \circ_\kappa$ and $\|z - \xi\| \leq \Delta$, and if $\beta \in V \cap \Gamma(z)$, then

$$(6.3) \quad |L(z, \beta) - L(\xi, \beta)| \leq \gamma.$$

Since φ is piecewise linear, for all but finitely many $t \in [0, 1] \setminus \{t^{(0)}, t^{(1)}, \dots, t^{(N)}\}$ we have $\dot{\varphi}(t) \in \Gamma(\varphi(t))$. By adding additional points $t^{(i)}$ if necessary, we can assume without loss of generality that the following three items hold.

- For each $i \in \{0, 1, \dots, N - 1\}$ there is a sequence of symbols $(\bar{\circ}_1, \dots, \bar{\circ}_\kappa)$ such that $\varphi(t) \in \mathcal{F}^{(\bar{\circ}_1, \dots, \bar{\circ}_\kappa)}$ for all $t \in (t^{(i)}, t^{(i+1)})$.
- For each $i \in \{0, 1, \dots, N - 1\}$ and any point $t \in (t^{(i)}, t^{(i+1)})$ $\dot{\varphi}(t) \in \Gamma(\varphi(t))$.
- For each $i \in \{0, 1, \dots, N - 1\}$

$$(6.4) \quad \sup \left\{ \|\varphi(t) - \varphi(t_i)\| : t \in [t^{(i)}, t^{(i+1)}] \right\} \leq \Delta/4.$$

Now consider any pair of points z and ξ both lying in the same set $V \cap \mathcal{F}^{(\circ_1, \dots, \circ_\kappa)}$ for some choice of $\circ_1, \dots, \circ_\kappa$ and take $v \in \Sigma(\xi)$. Then $r(\xi, v) > 0$ and since $\Sigma(\xi) = \Sigma(z)$, we also have $r(z, v) > 0$. By formula (4.1) there exists a constant $C \in [1, \infty)$ with the property that

$$1/C \leq r(\xi, v) \leq C \text{ and } 1/C \leq r(z, v) \leq C.$$

Since $r(\cdot, v)$ is uniformly continuous on each set $V \cap \mathcal{F}^{(\circ_1, \dots, \circ_\kappa)}$, it follows from Lemma 3.9 that there exists $\Delta_2 \in (0, \Delta_1)$ with the following property. If $u \in [0, \infty)$, $\Delta \in (0, \Delta_2)$, z and ξ both lie in the same set $V \cap \mathcal{F}^{(\circ_1, \dots, \circ_\kappa)}$ with $\|z - \xi\| \leq \Delta$, and $v \in \Sigma(\xi)$, then

$$\begin{aligned} (1 - \gamma)r(z, v)\ell(u/r(z, v)) - \gamma \\ \leq r(\xi, v)\ell(u/r(\xi, v)) \leq (1 + \gamma)r(z, v)\ell(u/r(z, v)) + \gamma. \end{aligned}$$

Finally, let \tilde{x} be any point in \mathcal{S}^∞ and assume that ξ is a point in \mathbb{R}^d lying in a facet $\mathcal{F}^{(\circ_1, \dots, \circ_\kappa)}$ that either contains \tilde{x} or is adjacent to \tilde{x} . The point ξ also lies in a facet $\tilde{\mathcal{F}}^{(\circ_1, \dots, \circ_\kappa(\tilde{x}))}$ determined by \tilde{x} . By letting z in the last display converge to \tilde{x} with $z - \tilde{x}$ always lying in the same facet $\tilde{\mathcal{F}}^{(\circ_1, \dots, \circ_\kappa(\tilde{x}))}$ as ξ , we conclude from the definition (6.1) of $\tilde{r}(\tilde{x}; \cdot, \cdot)$ that if $u \in [0, \infty)$, $\Delta \in (0, \Delta_2)$, $\|\xi - \tilde{x}\| \leq \Delta/2$, and $v \in \Sigma(\xi)$, then

$$(6.5) \quad \begin{aligned} (1 - \gamma)\tilde{r}(\tilde{x}; \xi, v)\ell(u/\tilde{r}(\tilde{x}; \xi, v)) - \gamma \\ \leq r(\xi, v)\ell(u/r(\xi, v)) \leq (1 + \gamma)\tilde{r}(\tilde{x}; \xi, v)\ell(u/\tilde{r}(\tilde{x}; \xi, v)) + \gamma. \end{aligned}$$

We now proceed with the proof of Proposition 5.1. For each $i \in \{0, 1, \dots, N - 1\}$ we set $z^{(i)} \doteq \varphi(t^{(i)})$, and for some $\tilde{t} \in (t^{(i)}, t^{(i+1)})$ we define $\tilde{x}^{(i)} \doteq \varphi(\tilde{t})$ and $\beta^{(i)} \doteq \dot{\varphi}(\tilde{t})$. The definition of $\beta^{(i)}$ is independent of the choice of \tilde{t} in

this interval. For $s \in (t^{(i)}, t^{(i+1)})$ we have

$$\varphi(s) = \varphi(t^{(i)}) + (s - t^{(i)})\dot{\varphi}(\tilde{t}) = z^{(i)} + (s - t^{(i)})\beta^{(i)}.$$

For $n \in \mathbb{N}$ and $w \in \mathcal{S}^n$ let $|\Sigma(w)|$ denote the number of elements in the set $\Sigma(w)$. Condition 4.1 implies that there exists $M \in \mathbb{N}$ so that

$$(6.6) \quad \sup_{n \in \mathbb{N}} \sup_{w \in \mathcal{S}^n} |\Sigma(w)| \leq M.$$

Suppose we prove that

$$(6.7) \quad \lim_{\varepsilon \rightarrow 0} \lim_{\delta \rightarrow 0} \liminf_{n \rightarrow \infty} \inf_{\{y \in \mathcal{S}^n : \|y - z^{(i)}\| \leq \delta\}} \left(-\frac{1}{n} \log P_{y, t^{(i)}}^n \left\{ \sup_{s \in [t^{(i)}, t^{(i+1)})} \|X^n(s) - z^{(i)} - (s - t^{(i)})\beta^{(i)}\| < \varepsilon \right\} \right) \geq (t^{(i+1)} - t^{(i)}) \cdot \left[(1 - \gamma) L(\tilde{x}^{(i)}, \beta^{(i)}) - M\gamma \right]$$

and

$$(6.8) \quad \lim_{\varepsilon \rightarrow 0} \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\{y \in \mathcal{S}^n : \|y - z^{(i)}\| \leq \delta\}} \left(-\frac{1}{n} \log P_{y, t_i}^n \left\{ \sup_{s \in [t^{(i)}, t^{(i+1)})} \|X^n(s) - z^{(i)} - (s - t^{(i)})\beta^{(i)}\| < \varepsilon \right\} \right) \leq (t^{(i+1)} - t^{(i)}) \cdot \left[(1 + \gamma) L(\tilde{x}^{(i)}, \beta^{(i)}) + M\gamma \right],$$

where $P_{y, t^{(i)}}^n$ denotes probability conditioned on $X^n(t^{(i)}) = y$. Then by using formulas (6.3) and (6.4) and by using the Markov property to combine the estimates over the intervals $(t^i, t^{(i+1)})$, $i \in \{0, 1, \dots, N - 1\}$, we obtain

$$\begin{aligned} & \lim_{\varepsilon \rightarrow 0} \lim_{\delta \rightarrow 0} \liminf_{n \rightarrow \infty} \inf_{\{y \in \mathcal{S}^n : \|y - x\| \leq \delta\}} \left(-\frac{1}{n} \log P_{y, 0}^n \left\{ \sup_{s \in [0, 1]} \|X^n(s) - \varphi(s)\| < \varepsilon \right\} \right) \\ & \geq (1 - \gamma) \int_0^1 L(\varphi(s), \dot{\varphi}(s)) ds - (1 - \gamma)\gamma - M\gamma \\ & \geq (1 - \gamma) I_x(\varphi) - (M + 1)\gamma \end{aligned}$$

and

$$\begin{aligned} & \lim_{\varepsilon \rightarrow 0} \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\{y \in \mathcal{S}^n : \|y - x\| \leq \delta\}} \left(-\frac{1}{n} \log P_{y, 0}^n \left\{ \sup_{s \in [0, 1]} \|X^n(s) - \varphi(s)\| < \varepsilon \right\} \right) \\ & \leq (1 + \gamma) \int_0^1 L(\varphi(s), \dot{\varphi}(s)) ds + (1 + \gamma)\gamma + M\gamma \\ & = (1 + \gamma) I_x(\varphi) + (M + 1)\gamma + \gamma^2. \end{aligned}$$

Since $\gamma > 0$ is arbitrary, we will have established Proposition 5.1. \square

Thus it remains to prove the limits (6.7) and (6.8). In order to simplify the notation we will write $0, z, \tilde{x}, \beta,$ and τ in place of $t^{(i)}, z^{(i)}, \tilde{x}^{(i)}, \beta^{(i)},$ and $t^{(i+1)} - t^{(i)},$ respectively. Thus $z = \varphi(0), \tilde{x} = \varphi(\tilde{t}),$ and $\beta = \dot{\varphi}(\tilde{t}),$ where \tilde{t} is some point in the interval $(0, \tau).$ For all $s \in (0, \tau)$ the path $\varphi(s) = z + s\beta$ lies in some fixed facet $\mathcal{F}^{(\sigma_1, \dots, \sigma_k)}.$ With this new notation equation (6.7) becomes

$$(6.9) \quad \lim_{\varepsilon \rightarrow 0} \lim_{\delta \rightarrow 0} \liminf_{n \rightarrow \infty} \inf_{\{y \in \mathcal{S}^n : \|y - z\| \leq \delta\}} \left(-\frac{1}{n} \log P_{y,0}^n \left\{ \sup_{s \in [0, \tau]} \|X^n(s) - z - s\beta\| < \varepsilon \right\} \right) \geq \tau [(1 - \gamma) L(\tilde{x}, \beta) - M\gamma]$$

and equation (6.8) becomes

$$(6.10) \quad \lim_{\varepsilon \rightarrow 0} \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\{y \in \mathcal{S}^n : \|y - z\| \leq \delta\}} \left(-\frac{1}{n} \log P_{y,0}^n \left\{ \sup_{s \in [0, \tau]} \|X^n(s) - z - s\beta\| < \varepsilon \right\} \right) \leq \tau [(1 + \gamma) L(\tilde{x}, \beta) + M\gamma].$$

Suppose that we were able to replace $X^n(s) - z - s\beta$ in formulas (6.9) and (6.10) by $\tilde{X}_{\tilde{x}}^n(s) - z - s\beta.$ In that case the limits (6.9) and (6.10) would follow immediately from the limits in formula (4.7). In order to apply these limits, we must check that

$$(6.11) \quad z \in \Gamma(\tilde{x}) \text{ and } \beta \in \Gamma(\tilde{x}).$$

Clearly $\tilde{x} \in \Gamma(\tilde{x}),$ and by the choice of the number \tilde{t} we have $\beta \in \Gamma(\tilde{x}).$ Since $\varphi(\tilde{t}) = z + \tilde{t}\beta,$ it follows that $z \in \Gamma(\tilde{x}).$ The technique that we will use to prove the limits (6.9) and (6.10) will be to compare the representation formulas for the probability appearing in these equations with those for the probability appearing in formula (4.7). There are two sources of error with which we must contend. The first source of error is the variation of the rates $r(\xi, v)$ as a function of ξ within each facet. This has been dealt with by taking $\Delta > 0$ sufficiently small [see equations (6.3) and (6.5)]. The second source of error is due to the fact that $\varphi(s)$ lies in the facet $\mathcal{F}^{(\sigma_1, \dots, \sigma_k)}$ only for s in the open time interval $(0, \tau)$ and not necessarily for $s = 0$ or $s = \tau.$ In fact for $s = 0$ or $s = \tau$ $\varphi(s)$ may lie in a different facet. As a result, the process $X^n(s) - z - s\beta$ may have “boundary effects” that the process $\tilde{X}_{\tilde{x}}^n(s) - z - s\beta,$ defined in terms of the local jump intensity function $\tilde{r}(\tilde{x}; \cdot, \cdot),$ will not have. A comparison of the two processes may therefore not be useful for times near the end points 0 and $\tau.$ We will show that the error due to this discrepancy can be made as small as desired by using the communication/controllability condition expressed in Condition 3.4.

In order to prove the lower bound (6.9) and the upper bound (6.10), we must adapt to the processes mentioned in the last paragraph the stochastic control representation formula given in Theorem 3.3. The adaptations are straightforward. The only difference is that we will replace the interval $[0, 1]$ by the interval

$[t_1, t_2]$, where t_1 and t_2 are numbers satisfying $0 \leq t_1 \leq t_2 \leq 1$. For completeness we will give a full description. Theorem 3.3 was stated in Section 3, in which Condition 3.1 requires the jump intensity function to be constant on each facet. However, this condition is not needed in the proof of the representation formula. In fact it holds for any jump Markov process satisfying Condition 4.1.

Let $\{X(t), t \in [0, 1]\}$ be a jump Markov process with state space \mathcal{S} and jump intensity function $r(\cdot, \cdot)$ satisfying Condition 4.1. For $n \in \mathbb{N}$, $y \in \mathcal{S}^n$, numbers t_1 and t_2 satisfying $0 \leq t_1 \leq t_2 \leq 1$, $z \in \mathcal{S}^\infty$, $\beta \in \mathbb{R}^d$, and $\varepsilon > 0$, we consider the scaled jump Markov process $\{X^n(t), t \in [0, 1]\}$ with state space \mathcal{S}^n and in terms of it define the quantity

$$q^n(y, t_1, t_2; z, \beta, \varepsilon) \doteq -\frac{1}{n} \log P_{y, t_1}^n \left\{ \sup_{s \in [t_1, t_2]} \|X^n(s) - z - (s - t_1)\beta\| < \varepsilon \right\}.$$

A measurable function $u^n(y, v, t)$ mapping $\mathcal{S}^n \times \mathbb{Z}^d \times [0, 1]$ into $[0, \infty)$ is said to be a control. We will take as the class of admissible controls the same class characterized by Definition 3.2 and used in Section 3. Given an admissible control, let $\{\Xi^n(t), t \in [t_1, t_2]\}$ be a controlled process associated with this control as in Definition 3.2; we make the obvious modification of replacing the interval $[0, 1]$ by the interval $[t_1, t_2]$. Such a controlled process is a nonstationary jump Markov process having the generators \mathcal{L}^n defined in formula (3.16). In order to specify the cost structure, we define for $\varphi \in \mathcal{D}([t_1, t_2]: \mathbb{R}^d)$ the exit time

$$\pi(t_1, t_2, \varphi; z, \beta, \varepsilon) \doteq \inf\{s \in [t_1, t_2]: \|\varphi(s) - z - (s - t_1)\beta\| \geq \varepsilon\}$$

and follow the convention that the infimum over the empty set is ∞ . Thus $\pi(t_1, t_2, \varphi; z, \beta, \varepsilon) = \infty$ if and only if $\|\varphi(s) - z - (s - t_1)\beta\| < \varepsilon$ for all $s \in [t_1, t_2]$. For $w \in [0, 1] \cup \{\infty\}$ we also define the function

$$g(w) \doteq \begin{cases} 0 & \text{if } w = \infty, \\ \infty & \text{if } w \in [0, 1]. \end{cases}$$

The definition of the stochastic control problem is completed by defining the minimal cost function

(6.12)

$$V^n(y, t_1, t_2; z, \beta, \varepsilon) \doteq \inf_{E_{y, t_1}^n} \left\{ \int_{t_1}^{\pi(t_1, t_2, \Xi^n; x, \beta, \varepsilon) \wedge t_2} \sum_{v \in \mathbb{Z}^d} r(\Xi^n(s), v) \ell \left(\frac{u^n(\Xi^n(s), v, s)}{r(\Xi^n(s), v)} \right) ds + g(\pi(t_1, t_2, \Xi^n; z, \beta, \varepsilon)) \right\}.$$

The infimum in this formula is taken over all admissible controls u^n , $E_{y, t_1}^n\{\cdot\}$ denotes expectation conditioned on $\Xi^n(t_1) = y$, and ℓ is the nonnegative function defined in formula (3.18). Exactly as in Theorem 3.3 we can prove, with

only notational changes, the representation

$$(6.13) \quad q^n(y, t_1, t_2; z, \beta, \varepsilon) = V^n(y, t_1, t_2; z, \beta, \varepsilon).$$

Although Theorem 3.3 was stated with $t_2 = 1$, the introduction of the extra parameter $t_2 \in [0, 1]$ causes no difficulty.

Fixing \tilde{x} in the limit state space \mathcal{S}^∞ , we must also consider the analogous quantities for the sequence of scaled jump Markov processes $\{\tilde{X}_{\tilde{x}}^n(t), t \in [0, 1]\}$ arising from the local jump intensity function $\tilde{r}(\tilde{x}; y, v)$ defined in formula (6.1). These processes have the respective state spaces $\{\mathcal{S}_{\tilde{x}}^n, n \in \mathbb{N}\}$. For $n \in \mathbb{N}$, $y \in \mathcal{S}_{\tilde{x}}^n$, numbers t_1 and t_2 satisfying $0 \leq t_1 \leq t_2 \leq 1$, $z \in \mathcal{S}^\infty$, $\beta \in \mathbb{R}^d$, and $\varepsilon > 0$, we define the quantity

$$\tilde{q}_{\tilde{x}}^n(y, t_1, t_2; z, \beta, \varepsilon) \doteq -\frac{1}{n} \log P_{\tilde{x}; y, t_1}^n \left\{ \sup_{s \in [t_1, t_2]} \|\tilde{X}_{\tilde{x}}^n(s) - z - (s - t_1)\beta\| < \varepsilon \right\}.$$

For points $z \in \Gamma(\tilde{x})$ and $\beta \in \Gamma(\tilde{x})$ these quantities satisfy the limits in formula (4.7). The definition of the associated stochastic control problem is given by

$$(6.14)$$

$$\begin{aligned} &\tilde{V}_{\tilde{x}}^n(y, t_1, t_2; z, \beta, \varepsilon) \\ &\doteq \inf_{\tilde{E}_{\tilde{x}; y, t_1}^n} \left\{ \int_{t_1}^{\pi(t_1, t_2, \tilde{\Xi}_{\tilde{x}}^n; z, \beta, \varepsilon) \wedge t_2} \sum_{v \in \mathbb{Z}^d} \tilde{r}(\tilde{x}; \tilde{\Xi}_{\tilde{x}}^n(s), v) \ell \left(\frac{u^n(\tilde{\Xi}_{\tilde{x}}^n(s), v, s)}{\tilde{r}(\tilde{x}; \tilde{\Xi}_{\tilde{x}}^n(s), v)} \right) ds \right. \\ &\quad \left. + g(\pi(t_1, t_2, \tilde{\Xi}_{\tilde{x}}^n; z, \beta, \varepsilon)) \right\}, \end{aligned}$$

where the associated controlled process $\tilde{\Xi}_{\tilde{x}}^n$ has the state space $\mathcal{S}_{\tilde{x}}^n$ and otherwise is defined in exactly the same way as the controlled process Ξ^n appearing in formula (6.12). The infimum in equation (6.14) is taken over all admissible controls u^n and $\tilde{E}_{\tilde{x}; y, t_1}^n$ denotes expectation conditioned on $\tilde{\Xi}_{\tilde{x}}^n(t_1) = y$. All the other quantities are as in the last paragraph. Exactly as in Theorem 3.3 we can prove, with only notational changes, the representation

$$(6.15) \quad \tilde{q}_{\tilde{x}}^n(y, t_1, t_2; z, \beta, \varepsilon) = \tilde{V}_{\tilde{x}}^n(y, t_1, t_2; z, \beta, \varepsilon).$$

In the proof of the large deviation limits in equations (6.9) and (6.10), we will relate the original controlled process Ξ^n and the localized controlled process $\tilde{\Xi}_{\tilde{x}}^n$ and we will use inequality (6.5) relating the original jump intensity function $r(\cdot, \cdot)$ and the localized jump intensity function $\tilde{r}(\tilde{x}; \cdot, \cdot)$. We are now ready to prove these large deviation limits.

Proof of the lower bound (6.9). We recall the numbers Δ_2 and $\Delta \in (0, \Delta_2)$, which arose in the discussion leading up to formula (6.5), as well as the number τ , which was introduced just before the lower bound (6.9). We also recall that for the given function φ we have set $z \doteq \varphi(0)$, $\tilde{x} \doteq \varphi(\tilde{i})$, and $\beta \doteq \dot{\varphi}(\tilde{i})$, where

\tilde{t} is some point in the interval $(0, \tau)$. For all $s \in (0, \tau)$ the path $\varphi(s) = z + s\beta$ lies in some fixed facet $\mathcal{F}^{(\delta_1, \dots, \delta_k)}$.

Let η be a parameter in the interval $(0, \tau/2)$. We can choose $\mu \in (0, \Delta/4)$ so that the following holds.

- Whenever $w \in \mathbb{R}^d$ satisfies $\|w - \varphi(s)\| < \mu$ for some $s \in [\eta, \tau - \eta]$, then w lies in the same facet $\mathcal{F}^{(\delta_1, \dots, \delta_k)}$ as $\varphi(s)$ or in an adjacent facet; namely, in a facet that does not contain $\varphi(s)$ but whose closure contains $\varphi(s)$.

It follows that if $\{\xi(s), s \in [\eta, \tau - \eta]\}$ is any function in $\mathcal{D}([\eta, \tau - \eta] : \mathcal{S}^n)$ satisfying

$$(6.16) \quad \|\xi(s) - \varphi(s)\| < \mu \text{ for all } s \in [\eta, \tau - \eta],$$

then for all $n \in \mathbb{N}$

$$(6.17) \quad \bigcup_{s \in [\eta, \tau - \eta]} \{\xi(s)\} \subset \mathcal{S}^n \cap \mathcal{S}_x^n.$$

Given $\alpha \in (0, \mu)$, the communication/controllability condition expressed in Condition 3.4 guarantees the existence of an integer $N_1 \in \mathbb{N}$ and a number $c > 0$, both depending on α , such that the quantity

$$(6.18) \quad q^n(y, 0, \tau; z, \beta, \alpha) \doteq -\frac{1}{n} \log P_{y,0}^n \left\{ \sup_{s \in [0, \tau]} \|X^n(s) - z - s\beta\| < \alpha \right\}$$

is finite for all $n \in \mathbb{N}$ satisfying $n \geq N_1$ and all $y \in \mathcal{S}^n$ satisfying $\|y - z\| \leq c$. For the remainder of the proof of the lower bound (6.9), we work with such values of n and y .

We now choose \bar{u}^n to be an admissible control that is η -optimal in the representation (6.13) for $q^n(y, 0, \tau; z, \beta, \alpha)$. In other words, if $\{\Xi^n(t), t \in [0, \tau]\}$ is the associated controlled process, then $\Xi^n(0) = y$ and the cost with respect to \bar{u}^n is within η of the infimum $V^n(y, 0, \tau; z, \beta, \varepsilon)$. This implies that

$$\begin{aligned} \infty > q^n(y, 0, \tau; z, \beta, \alpha) &= V^n(y, 0, \tau; z, \beta, \alpha) \\ &\geq \bar{E}_{y,0}^n \{g(\pi(0, \tau, \Xi^n; z, \beta, \alpha))\} - \eta, \end{aligned}$$

which in turn implies that with probability 1

$$g(\pi(0, \tau, \Xi^n; z, \beta, \alpha)) = 0 \text{ and thus } \pi(0, \tau, \Xi^n; z, \beta, \alpha) = \infty.$$

Hence with probability 1

$$(6.19) \quad \|\Xi^n(s) - z - s\beta\| = \|\Xi^n(s) - \varphi(s)\| < \alpha \text{ for all } s \in [0, \tau].$$

In particular, the value $\xi_0 \doteq \Xi^n(\eta)$ satisfies the inequality $\|\xi_0 - z - \eta\beta\| < \alpha$. Thus for any $n \in \mathbb{N}$ satisfying $n \geq N_1$ and for any $y \in \mathcal{S}^n$ satisfying

$\|y - z\| \leq c$ we have

(6.20)

$$\begin{aligned} q^n(y, 0, \tau; z, \beta, \alpha) &= V^n(y, 0, \tau; z, \beta, \alpha) \\ &\geq E_{y,0}^n \left\{ \int_0^\tau \sum_{v \in \mathbb{Z}^d} r(\Xi^n(s), v) \ell \left(\frac{\bar{u}^n(\Xi^n(s), v, s)}{r(\Xi^n(s), v)} \right) ds \right\} - \eta. \\ &\geq E_{y,0}^n \left\{ \int_\eta^{\tau-\eta} \sum_{v \in \mathbb{Z}^d} r(\Xi^n(s), v) \ell \left(\frac{\bar{u}^n(\Xi^n(s), v, s)}{r(\Xi^n(s), v)} \right) ds \right\} - \eta \\ &\geq \inf_{\{\xi_0 \in \mathcal{S}^n : \|\xi_0 - z - \eta\beta\| < \alpha\}} E_{\xi_0, \eta}^n \left\{ \int_\eta^{\tau-\eta} \sum_{v \in \mathbb{Z}^d} r(\Xi^n(s), v) \ell \left(\frac{\bar{u}^n(\Xi^n(s), v, s)}{r(\Xi^n(s), v)} \right) ds \right\} \\ &\quad - \eta. \end{aligned}$$

For all $s \in (0, \tau)$ $\varphi(s)$, and thus $\tilde{x} \doteq \varphi(\tilde{t})$, lie in the facet $\mathcal{F}^{(\sigma_1, \dots, \sigma_k)}$. Inequality (6.19), the definition of $\mu \in (0, \Delta/4)$, and the selection of $\alpha \in (0, \mu)$ imply that for all $s \in [\eta, \tau - \eta]$ $\Xi^n(s)$ lies in the same facet $\mathcal{F}^{(\sigma_1, \dots, \sigma_k)}$ or in an adjacent facet. In addition, using formulas (6.4) and (6.19), we see that for all $s \in [\eta, \tau - \eta]$

$$\|\Xi^n(s) - \tilde{x}\| \leq \|\Xi^n(s) - \varphi(s)\| + \|\varphi(s) - \varphi(\tilde{t})\| < \alpha + \Delta/4 < \mu + \Delta/4 \leq \Delta/2.$$

Hence we can apply the first inequality in formula (6.5). For any $n \in \mathbb{N}$ satisfying $n \geq N_1$ and for any $\xi_0 \in \mathcal{S}^n$ satisfying $\|\xi_0 - z - \eta\beta\| < \alpha$, we have from the property of the constant M expressed in formula (6.6) that

(6.21)

$$\begin{aligned} E_{\xi_0, \eta}^n &\left\{ \int_\eta^{\tau-\eta} \sum_{v \in \mathbb{Z}^d} r(\Xi^n(s), v) \ell \left(\frac{\tilde{u}^n(\Xi^n(s), v, s)}{r(\Xi^n(s), v)} \right) ds \right\} \\ &\geq (1 - \gamma) \cdot E_{\xi_0, \eta}^n \left\{ \int_\eta^{\tau-\eta} \sum_{v \in \mathbb{Z}^d} \tilde{r}(\tilde{x}; \Xi^n(s), v) \ell \left(\frac{u^n(\Xi^n(s), v, s)}{\tilde{r}(\tilde{x}; \Xi^n(s), v)} \right) ds \right\} \\ &\quad - M\gamma\tau. \end{aligned}$$

Combining this with display (6.20), we see that for any $n \in \mathbb{N}$ satisfying $n \geq N_1$

(6.22)

$$\begin{aligned} &\inf_{\{y \in \mathcal{S}^n : \|y - z\| \leq c\}} \left(-\frac{1}{n} \log P_{y,0} \left\{ \sup_{s \in [0, \tau]} \|X^n(s) - z - s\beta\| < \alpha \right\} \right) \\ &\geq (1 - \gamma) \inf_{\{\xi_0 \in \mathcal{S}^n : \|\xi_0 - z - \eta\beta\| < \alpha\}} E_{\xi_0, \eta}^n \left\{ \int_\eta^{\tau-\eta} \sum_{v \in \mathbb{Z}^d} \tilde{r}(\tilde{x}; \Xi^n(s), v) \ell \left(\frac{\bar{u}^n(\Xi^n(s), v, s)}{\tilde{r}(\tilde{x}; \Xi^n(s), v)} \right) ds \right\} - \eta - M\gamma\tau. \end{aligned}$$

We claim that the control \bar{u}^n and the associated controlled process $\{\Xi^n(s), s \in [\eta, \tau - \eta]\}$ are also an admissible control and associated controlled process to apply in the control representation (6.15) for $\tilde{q}_{\tilde{x}}(\xi_0, \eta, \tau - \eta; \cdot, \cdot, \cdot)$. This follows once we show that

$$\bigcup_{s \in [\eta, \tau - \eta]} \{\Xi^n(s)\} \subset \mathcal{S}^n \cap \mathcal{S}_{\tilde{x}}^n;$$

in other words, the points visited by $\{\Xi^n(s), s \in [\eta, \tau - \eta]\}$ are common to the two state spaces \mathcal{S}^n and $\mathcal{S}_{\tilde{x}}^n$. The last display is a consequence of inequality (6.19), the definition of μ , the selection of $\alpha \in (0, \mu)$, and the fact that formula (6.17) holds for any function $\xi \in \mathcal{D}([\eta, \tau - \eta]: \mathcal{S}^n)$ satisfying equation (6.16). Inequality (6.19) also implies that with probability 1

$$\|\Xi^n(s) - z - s\beta\| = \|\Xi^n(s) - (z + \eta\beta) - (s - \eta)\beta\| < \alpha \text{ for all } s \in [\eta, \tau - \eta],$$

which in turn implies that with probability 1

$$\pi(\eta, \tau - \eta, \Xi^n; z + \eta\beta, \beta, \alpha) = \infty$$

and thus

$$g(\pi(\eta, \tau - \eta, \Xi^n; z + \eta\beta, \beta, \alpha)) = 0.$$

We now use the representation formula (6.15) and formula (6.22) to say that for any $n \in \mathbb{N}$ satisfying $n \geq N_1$

(6.23)

$$\begin{aligned} & \inf_{\{y \in \mathcal{S}^n : \|y - z\| \leq c\}} \left(-\frac{1}{n} \log P_{y,0}^n \left\{ \sup_{s \in [0, \tau]} \|X^n(s) - z - s\beta\| < \alpha \right\} \right) \\ & \geq (1 - \gamma) \inf_{\{\xi_0 \in \mathcal{S}^n : \|\xi_0 - z - \eta\beta\| < \alpha\}} \\ & \quad E_{\xi_0, \eta}^n \left\{ \int_{\eta}^{\tau - \eta} \sum_{v \in \mathbb{Z}^d} \tilde{r}(\tilde{x}; \Xi^n(s), v) \ell \left(\frac{\bar{u}^n(\Xi^n(s), v, s)}{\tilde{r}(\tilde{x}; \Xi^n(s), v)} \right) ds \right\} \log P^n - \eta - M\gamma\tau \\ & = (1 - \gamma) \inf_{\{\xi_0 \in \mathcal{S}^n : \|\xi_0 - z - \eta\beta\| < \alpha\}} \\ & \quad E_{\xi_0, \eta}^n \left\{ \int_{\eta}^{\tau - \eta} \sum_{v \in \mathbb{Z}^d} \tilde{r}(\tilde{x}; \Xi^n(s), v) \ell \left(\frac{\bar{u}^n(\Xi^n(s), v, s)}{\tilde{r}(\tilde{x}; \Xi^n(s), v)} \right) ds \right. \\ & \quad \left. + g(\pi(\eta, \tau - \eta, \Xi^n; z + \eta\beta, \beta, \alpha)) \right\} - \eta - M\gamma\tau \\ & \geq (1 - \gamma) \cdot \inf_{\{\xi_0 \in \mathcal{S}^n : \|\xi_0 - z - \eta\beta\| < \alpha\}} \tilde{V}_{\tilde{x}}^n(\xi_0, \eta, \tau - \eta; z + \eta\beta, \beta, \alpha) - \eta - M\gamma\tau \\ & = (1 - \gamma) \cdot \inf_{\{\xi_0 \in \mathcal{S}^n : \|\xi_0 - z - \eta\beta\| < \alpha\}} \tilde{q}_{\tilde{x}}^n(\xi_0, \eta, \tau - \eta; z + \eta\beta, \beta, \alpha) - \eta - M\gamma\tau. \end{aligned}$$

Formula (6.11) implies that the points $z + \eta\beta$ and β both lie in the set $\Gamma(\tilde{x})$. Hence we can apply the first limit in equation (4.7), which yields

$$\begin{aligned} & \lim_{\alpha \rightarrow 0} \liminf_{n \rightarrow \infty} \inf_{\{\xi_0 \in \mathcal{S}_{\tilde{x}}^n : \|\xi_0 - z - \eta\beta\| < \alpha\}} \tilde{q}_{\tilde{x}}^n(\xi_0, \eta, \tau - \eta; z + \eta\beta, \beta, \alpha) \\ &= \lim_{\alpha \rightarrow 0} \liminf_{n \rightarrow \infty} \inf_{\{\xi_0 \in \mathcal{S}_{\tilde{x}}^n : \|\xi_0 - z - \eta\beta\| < \alpha\}} \\ & \quad \left(-\frac{1}{n} \log P_{\tilde{x}; \xi_0, \eta}^n \left\{ \sup_{s \in [\eta, \tau - \eta]} \|\tilde{X}_{\tilde{x}}^n(s) - (z + \eta\beta) - (s - \eta)\beta\| < \alpha \right\} \right) \\ & \geq \lim_{\varepsilon \rightarrow 0} \lim_{\alpha \rightarrow 0} \liminf_{n \rightarrow \infty} \inf_{\{\xi_0 \in \mathcal{S}_{\tilde{x}}^n : \|\xi_0 - z - \eta\beta\| < \alpha\}} \\ & \quad \left(-\frac{1}{n} \log P_{\tilde{x}; \xi_0, \eta}^n \left\{ \sup_{s \in [\eta, \tau - \eta]} \|\tilde{X}_{\tilde{x}}^n(s) - (z + \eta\beta) - (s - \eta)\beta\| < \varepsilon \right\} \right) \\ &= (\tau - 2\eta) \cdot L(\tilde{x}, \beta). \end{aligned}$$

Returning to display (6.23), we take the limit inferior as $n \rightarrow \infty$ and send $c \rightarrow 0$, then $\alpha \rightarrow 0$, and finally $\eta \rightarrow 0$ to conclude that

$$\begin{aligned} & \lim_{\alpha \rightarrow 0} \lim_{c \rightarrow 0} \liminf_{n \rightarrow \infty} \inf_{\{y \in \mathcal{S}^n : \|y - z\| \leq c\}} \left(-\frac{1}{n} \log P_{y, 0} \left\{ \sup_{s \in [0, \tau]} \|X^n(s) - z - s\beta\| < \alpha \right\} \right) \\ & \geq \tau [(1 - \gamma) L(\tilde{x}, \beta) - M\gamma]. \end{aligned}$$

This proves the lower bound (6.9). \square

Proof of the upper bound (6.10). In this proof we will have to adapt Lemma 3.5 to the current setting, in which Condition 3.1 is replaced by Condition 4.1. As noted in Remark 3.6, all the conclusions of Lemma 3.5 continue to hold provided we have the bounds (3.26) and (3.27). Under Conditions 3.4 and 4.1 these are valid [see equations (4.1) and (6.6)]. Because of this, in the sequel we will continue to refer to Lemma 3.5 when we actually mean the version of that lemma with Condition 4.1 replacing Condition 3.1.

Let $\eta \in (0, 1)$ be given and select $N \in \mathbb{N}$, $\delta_0 > 0$, and $\sigma > 0$ as in part (a) of Lemma 3.5. We also choose $\delta \in (0, \delta_0)$. The constants B and K that will appear later in this proof are the same as those in Lemma 3.5. We recall from part (a) of Lemma 3.5 that for all sufficiently small $\eta \in (0, 1)$ we can take $\delta_0 = \eta/(2K)$ and $\sigma = \iota\eta$ for some constant ι . Hence we can arrange that

$$(6.24) \quad \delta_0 < 3\Delta/16 \text{ and } \sigma < \tau/2$$

by reducing η if necessary. We write the time interval appearing in the upper bound (6.10) as

$$[0, \tau] = [0, \sigma] \cup [\sigma, \tau - \sigma] \cup [\tau - \sigma, \tau].$$

Each of the three subintervals will be dealt with separately.

We choose a point $t^* \in (0, \sigma)$ such that $\|\varphi(t^*) - z\| \leq \delta/3$. We also let $\{z^n, n \in \mathbb{N}\}$ be any sequence of points converging to $\varphi(t^*)$ such that each z^n lies in the set $\mathcal{F}^{(\delta_1, \dots, \delta_k)} \cap \mathcal{S}^n$. There exists a positive integer N_1 such that for all $n \geq N_1$ we have $\|z^n - \varphi(t^*)\| \leq \delta/3$ and thus

$$\|z^n - z\| \leq \|z^n - \varphi(t^*)\| + \|\varphi(t^*) - z\| \leq 2\delta/3.$$

Now let y be any point in \mathcal{S}^n satisfying $\|y - z\| \leq \delta/3$. Then for all $n \in \mathbb{N}$ satisfying $n \geq N_1$ we have $\|y - z^n\| \leq \delta$. Hence by part (a) of Lemma

3.5, for all $n \geq N \vee N_1$ we can find an admissible control u_1^n on the first time interval $[0, \sigma]$ having the following properties: with probability 1 u_1^n steers the associated controlled process Ξ^n from y at time 0 to z^n at time σ ; with probability 1 u_1^n keeps the controlled process Ξ^n a distance no greater than $BK\delta$ from z^n on the entire time interval $[0, \sigma]$; the expected integrated running cost over the interval $[0, \sigma]$ is no greater than η . If it happens that the point $z \doteq \varphi(0)$ lies in a facet different from $\mathcal{F}^{(\delta_1, \dots, \delta_\kappa)}$, then this construction “jogs” the controlled process away from this other facet at a cheap cost. If z lies in the same facet $\mathcal{F}^{(\delta_1, \dots, \delta_\kappa)}$ as $\varphi(s)$ for $s \in (0, \tau)$, then nothing is lost.

We need the next lemma in order to define the control and the associated controlled process on the middle time interval $[\sigma, \tau - \sigma]$.

Lemma 6.2. *There exists $\mu \in (0, \Delta/8)$ and $N_2 \in \mathbb{N}$ so that the following holds. Whenever $w \in \mathbb{R}^d$ satisfies*

$$\|w - z^n - (s - \sigma)\beta\| < \mu \text{ for some } s \in [\sigma, \tau - \sigma] \text{ and } n \geq N \vee N_1 \vee N_2,$$

then w lies in the same facet $\mathcal{F}^{(\delta_1, \dots, \delta_\kappa)}$ as $\varphi(s)$ or in an adjacent facet.

Proof. We can choose $\mu \in (0, \Delta/8)$ so that the following is true. Whenever $w \in \mathbb{R}^d$ satisfies $\|w - \varphi(s)\| < 2\mu$ for some $s \in [t^*, \tau - t^*]$, then w lies in the same facet $\mathcal{F}^{(\delta_1, \dots, \delta_\kappa)}$ as $\varphi(s)$ or in an adjacent facet. Now suppose that $w \in \mathbb{R}^d$ satisfies $\|w - z^n - (s - \sigma)\beta\| < \mu$ for some $s \in [\sigma, \tau - \sigma]$ and $n \in \mathbb{N}$. We will show the existence of $N_2 \in \mathbb{N}$ such that if $n \geq N \vee N_1 \vee N_2$, then with the same value of s

$$(6.25) \quad \|w - \varphi(t^* + s - \sigma)\| < 2\mu.$$

We have $t^* \in (0, \sigma]$ and thus $t^* \leq t^* + s - \sigma \leq \tau - \sigma \leq \tau - t^*$. Hence formula (6.25) will complete the proof of the lemma. Since $\varphi(t^* + s - \sigma) = \varphi(t^*) + (s - \sigma)\beta$, we have

$$\begin{aligned} & \|w - \varphi(t^* + s - \sigma)\| \\ & \leq \|w - z^n - (s - \sigma)\beta\| + \|z^n + (s - \sigma)\beta - \varphi(t^* + s - \sigma)\| \\ & < \mu + \|z^n - \varphi(t^*)\|. \end{aligned}$$

Because of the limit $\lim_{n \rightarrow \infty} z^n = \varphi(t^*)$, there exists $N_2 \in \mathbb{N}$ such that $\|z^n - \varphi(t^*)\| < \mu$ for all $n \geq N_2$. For all $n \geq N \vee N_1 \vee N_2$ we now obtain formula (6.25). \square

Let $\{\xi(s), s \in [\sigma, \tau - \sigma]\}$ be any function in $\mathcal{D}([\sigma, \tau - \sigma] : \mathcal{S}_x^n)$ satisfying

$$(6.26) \quad \|\xi(s) - z^n - (s - \sigma)\beta\| < \mu \text{ for all } s \in [\sigma, \tau - \sigma] \text{ and some } n \geq N \vee N_1 \vee N_2.$$

Then Lemma 6.2 implies that

$$(6.27) \quad \bigcup_{s \in [\sigma, \tau - \sigma]} \{\xi(s)\} \subset \mathcal{S}_x^n \cap \mathcal{S}^n.$$

For $\varepsilon \in (0, \mu)$ we will next use the control representation (6.15) for $\tilde{q}_x^n(z^n, \sigma, \tau - \sigma; z^n, \beta, \varepsilon)$ to design an admissible control for the interval $[\sigma, \tau - \sigma]$.

This control will force the associated controlled process to stay a distance no greater than ε from the path $z^n + (s - \sigma)\beta$ for all $s \in [\sigma, \tau - \sigma]$.

The communication/controllability condition expressed in Condition 3.4 guarantees the existence of an integer $N_3 \in \mathbb{N}$ depending on $\varepsilon \in (0, \mu)$ such that the quantity

$$\begin{aligned} &\tilde{q}_{\tilde{x}}^n(z^n, \sigma, \tau - \sigma; z^n, \beta, \varepsilon) \\ &\doteq -\frac{1}{n} \log P_{\tilde{x}; z^n, \sigma}^n \left\{ \sup_{s \in [\sigma, \tau - \sigma]} \|\tilde{X}_{\tilde{x}}^n(s) - z^n - (s - \sigma)\beta\| < \varepsilon \right\} \end{aligned}$$

is finite for all $n \geq N_3$. For the remainder of the proof of the upper bound (6.10), we work with values of $n \geq N \vee N_1 \vee N_2 \vee N_3$. Now choose \tilde{u}^n to be an admissible control that is η -optimal in the representation (6.15) for $\tilde{q}_{\tilde{x}}^n(z^n, \sigma, \tau - \sigma; z^n, \beta, \varepsilon)$. In other words, if $\{\tilde{\Xi}_{\tilde{x}}^n(t), t \in [\sigma, \tau - \sigma]\}$ is the associated controlled process, then $\tilde{\Xi}_{\tilde{x}}^n(\sigma) = z^n$ and the cost with respect to \tilde{u}^n is within η of the infimum $\tilde{V}_{\tilde{x}}^n(z^n, \sigma, \tau - \sigma; z^n, \beta, \varepsilon)$. This implies that

$$\begin{aligned} &\tilde{E}_{\tilde{x}; z^n, \sigma}^n \{g(\pi(\sigma, \tau - \sigma, \tilde{\Xi}_{\tilde{x}}^n; z^n, \beta, \varepsilon))\} - \eta \\ &\leq \tilde{V}_{\tilde{x}}^n(z^n, \sigma, \tau - \sigma; z^n, \beta, \varepsilon) = \tilde{q}_{\tilde{x}}^n(z^n, \sigma, \tau - \sigma; z^n, \beta, \varepsilon) < \infty, \end{aligned}$$

which in turn implies that with probability 1

$$g(\pi(\sigma, \tau - \sigma, \tilde{\Xi}_{\tilde{x}}^n; z^n, \beta, \varepsilon)) = 0 \text{ and thus } \pi(\sigma, \tau - \sigma, \tilde{\Xi}_{\tilde{x}}^n; z^n, \beta, \varepsilon) = \infty.$$

Hence with probability 1 we have for all $n \geq N \vee N_1 \vee N_2 \vee N_3$

$$(6.28) \quad \|\tilde{\Xi}_{\tilde{x}}^n(s) - z^n - (s - \sigma)\beta\| < \varepsilon \text{ for all } s \in [\sigma, \tau - \sigma].$$

It follows that

$$\begin{aligned} (6.29) \quad &\tilde{E}_{\tilde{x}; z^n, \sigma}^n \left\{ \int_{\sigma}^{\tau - \sigma} \sum_{v \in \mathbb{Z}^d} \tilde{r}(\tilde{x}; \tilde{\Xi}_{\tilde{x}}^n(s), v) \ell \left(\frac{\tilde{u}^n(\tilde{\Xi}_{\tilde{x}}^n(s), v, s)}{\tilde{r}(\tilde{x}; \tilde{\Xi}_{\tilde{x}}^n(s), v)} \right) ds \right\} - \eta \\ &\leq \tilde{V}_{\tilde{x}}^n(z^n, \sigma, \tau - \sigma; z^n, \beta, \varepsilon) = \tilde{q}_{\tilde{x}}^n(z^n, \sigma, \tau - \sigma; z^n, \beta, \varepsilon). \end{aligned}$$

For all $s \in (0, \tau)$ $\varphi(s)$, and thus $\tilde{x} \doteq \varphi(\tilde{t})$, lie in the facet $\mathcal{F}^{(\sigma_1, \dots, \sigma_\kappa)}$. Lemma 6.2, inequality (6.28), and the selection of $\varepsilon \in (0, \mu)$ imply that for all $s \in [\sigma, \tau - \sigma]$ $\tilde{\Xi}_{\tilde{x}}^n(s)$ lies in the same facet $\mathcal{F}^{(\sigma_1, \dots, \sigma_\kappa)}$ or in an adjacent facet. Equation (6.24) yields that

$$\|z^n - z\| \leq 2\delta/3 < 2\delta_0/3 < \Delta/8,$$

and thus by formulas (6.4) and (6.28) we see that for all $s \in [\sigma, \tau - \sigma]$

$$\begin{aligned} \|\tilde{\Xi}_{\tilde{x}}^n(s) - \tilde{x}\| &\leq \|\tilde{\Xi}_{\tilde{x}}^n(s) - z^n - (s - \sigma)\beta\| + \|\varphi(\tilde{t}) - \varphi(s - \sigma)\| + \|z^n - z\| \\ &< \varepsilon + \Delta/4 + \Delta/8 < \mu + \Delta/4 + \Delta/8 \leq \Delta/2. \end{aligned}$$

Hence we can apply the second inequality in formula (6.5). For any $n \in \mathbb{N}$ satisfying $n \geq N \vee N_1 \vee N_2 \vee N_3$, we have from the property of the constant

M expressed in formula (6.6) that

(6.30)

$$\begin{aligned} & \tilde{E}_{\tilde{x}; z^n, \sigma}^n \left\{ \int_{\sigma}^{\tau-\sigma} \sum_{v \in \mathbb{Z}^d} r(\tilde{\Xi}_{\tilde{x}}^n(s), v) \ell \left(\frac{\tilde{u}^n(\tilde{\Xi}_{\tilde{x}}^n(s), v, s)}{r(\tilde{\Xi}_{\tilde{x}}^n(s), v)} \right) ds \right\} \\ & \leq (1 + \gamma) \cdot \tilde{E}_{\tilde{x}; z^n, \sigma}^n \left\{ \int_{\sigma}^{\tau-\sigma} \sum_{v \in \mathbb{Z}^d} \tilde{r}(\tilde{x}; \tilde{\Xi}_{\tilde{x}}^n(s), v) \ell \left(\frac{u^n(\tilde{\Xi}_{\tilde{x}}^n(s), v, s)}{\tilde{r}(\tilde{x}; \tilde{\Xi}_{\tilde{x}}^n(s), v)} \right) ds \right\} \\ & \quad + M\gamma\tau. \end{aligned}$$

Combining formulas (6.29) and (6.30) now yields that

(6.31)

$$\begin{aligned} & \tilde{E}_{\tilde{x}; z^n, \sigma}^n \left\{ \int_{\sigma}^{\tau-\sigma} \sum_{v \in \mathbb{Z}^d} r(\tilde{\Xi}_{\tilde{x}}^n(s), v) \ell \left(\tilde{u}^n(\tilde{\Xi}_{\tilde{x}}^n(s), v, s) / r(\tilde{\Xi}_{\tilde{x}}^n(s), v) \right) ds \right\} \\ & \leq (1 + \gamma) \cdot [\tilde{q}_{\tilde{x}}^n(z^n, \sigma, \tau - \sigma; z^n, \beta, \varepsilon) + \eta] + M\gamma\tau. \end{aligned}$$

The controlled process $\{\tilde{\Xi}_{\tilde{x}}^n(s), s \in [\sigma, \tau - \sigma]\}$ is associated with the control \tilde{u}^n . We claim that this control can be used to define an admissible control and an associated controlled process Ξ^n that can be applied in the control representation (6.13) for q^n . The claim follows once we show that

$$\bigcup_{s \in [\sigma, \tau - \sigma]} \{\tilde{\Xi}_{\tilde{x}}^n(s)\} \subset \mathcal{S}_{\tilde{x}}^n \cap \mathcal{S}^n;$$

in other words, the points visited by $\{\tilde{\Xi}_{\tilde{x}}^n(s), s \in [\sigma, \tau - \sigma]\}$ are common to the two state spaces $\mathcal{S}_{\tilde{x}}^n$ and \mathcal{S}^n . The last display is a consequence of inequality (6.28), the definition of μ , the selection of $\varepsilon \in (0, \mu)$, and the fact that formula (6.27) holds for any function $\xi \in \mathcal{D}([\sigma, \tau - \sigma] : \mathcal{S}_{\tilde{x}}^n)$ satisfying equation (6.26).

Finally we define the control on the final time interval $[\tau - \sigma, \tau]$. According to part (a) of Lemma 3.5, we can find an admissible control u_2^n on the time interval $[\tau - \sigma, \tau]$ having the following properties: with probability 1 u_2^n returns the controlled process Ξ^n at time τ to the position that it had at time $\tau - \sigma$; with probability 1 u_2^n keeps the controlled process on the entire time interval $[\tau - \sigma, \tau]$ a distance no greater than $BK\delta$ from the position that it had at time $\tau - \sigma$; the expected integrated running cost over the interval $[\tau - \sigma, \tau]$ is no greater than η .

Let $\alpha > 0$ be given and let y be any point in \mathcal{S}^n satisfying $\|y - z\| \leq \delta/3$. For $n \in \mathbb{N}$ satisfying $n \geq N \vee N_1 \vee N_2 \vee N_3$ we now define a composite control u^n on the time interval $[0, \tau]$ which will be applied in the control representation for

$$q^n(y, 0, \tau; z, \beta, \alpha) \doteq -\frac{1}{n} \log P_{y,0}^n \left\{ \sup_{s \in [0, \tau]} \|X^n(s) - z - s\beta\| < \alpha \right\}.$$

The control is given by the formula

$$u^n(\xi, v, t) = \begin{cases} u_1^n(\xi, v, t) & \text{for } t \in [0, \sigma], \\ \tilde{u}^n(\xi, v, t) & \text{for } t \in (\sigma, \tau - \sigma], \\ u_2^n(\xi, v, t) & \text{for } t \in (\tau - \sigma, \tau]. \end{cases}$$

Let $\{\Xi^n(s), s \in [0, \tau]\}$ denote the associated controlled process. We have the following properties.

- With probability 1, the controlled process Ξ^n satisfies $\Xi^n(0) = y$ and $\Xi^n(\sigma) = z^n$ and on the entire time interval $[0, \sigma]$ Ξ^n stays a distance no greater than $BK\delta$ from the point z^n . Since $\|z^n - z\| \leq 2\delta/3$, we have with probability 1

$$(6.32) \quad \sup_{s \in [0, \sigma]} \|\Xi^n(s) - z - s\beta\| \leq BK\delta + 2\delta/3 + \sigma\|\beta\|.$$

The expected integrated running cost over the interval $[0, \sigma]$ is no greater than η .

- The control \tilde{u}^n was used to define the controlled process $\tilde{\Xi}_x^n$ on the time interval $[\sigma, \tau - \sigma]$. With probability 1 $\tilde{\Xi}_x^n(\sigma) = z^n$. We are now using the same control \tilde{u}^n to define the associated controlled process Ξ^n on the time interval $[\sigma, \tau - \sigma]$. The definition of the control u_1^n implies that with probability 1 $\Xi^n(\sigma) = z^n$. It follows that the distribution of the controlled process $\{\Xi^n(s), s \in [\sigma, \tau - \sigma]\}$ with respect to $\tilde{E}_{z^n, \sigma}$ equals the distribution of the controlled process $\{\tilde{\Xi}_x^n(s), s \in [\sigma, \tau - \sigma]\}$ with respect to $\tilde{E}_{\tilde{x}; z^n, \sigma}$. Formula (6.28) then yields that with probability 1

$$\sup_{s \in [\sigma, \tau - \sigma]} \|\Xi^n(s) - z^n - (s - \sigma)\beta\| \leq \varepsilon.$$

Since $\|z^n - z\| \leq 2\delta/3$, it follows that with probability 1

$$(6.33) \quad \sup_{s \in [\sigma, \tau - \sigma]} \|\Xi^n(s) - z - s\beta\| \leq \varepsilon + 2\delta/3 + \sigma\|\beta\|.$$

In a moment [see equation (6.35)] we will estimate the expected integrated running cost over the interval $[\sigma, \tau - \sigma]$.

- With probability 1, on the entire time interval $[\tau - \sigma, \tau]$ the controlled process Ξ^n stays a distance no greater than $BK\delta$ from the position that it had at time $\tau - \sigma$. Equation (6.33) implies that with probability 1

$$(6.34) \quad \begin{aligned} & \sup_{s \in [\tau - \sigma, \tau]} \|\Xi^n(s) - z - s\beta\| \\ & \leq \sup_{s \in [\tau - \sigma, \tau]} \|\Xi^n(s) - \Xi^n(\tau - \sigma)\| + \|\Xi^n(\tau - \sigma) - z - (\tau - \sigma)\beta\| + \sigma\|\beta\| \\ & \leq BK\delta + [\varepsilon + 2\delta/3 + \sigma\|\beta\|] + \sigma\|\beta\|. \end{aligned}$$

The running cost over the interval $[\tau - \sigma, \tau]$ is no greater than η .

The distribution of the controlled process $\{\Xi^n(s), s \in [\sigma, \tau - \sigma]\}$ with respect to $\tilde{E}_{z^n, \sigma}$ equals the distribution of the controlled process $\{\tilde{\Xi}_x^n(s), s \in [\sigma, \tau - \sigma]\}$ with respect to $\tilde{E}_{\tilde{x}; z^n, \sigma}$. Hence formula (6.31) gives the following

estimate on the expected integrated running cost over the interval $[\sigma, \tau - \sigma]$:

(6.35)

$$\begin{aligned} & E_{z^n, \sigma}^n \left\{ \int_{\sigma}^{\tau - \sigma} \sum_{v \in \mathbb{Z}^d} r(\Xi^n(s), v) \ell(\tilde{u}^n(\Xi^n(s), v, s)/r(\Xi^n(s), v)) ds \right\} \\ &= \tilde{E}_{\tilde{x}^n; z^n, \sigma}^n \left\{ \int_{\sigma}^{\tau - \sigma} \sum_{v \in \mathbb{Z}^d} r(\tilde{\Xi}_{\tilde{x}^n}^n(s), v) \ell(\tilde{u}^n(\tilde{\Xi}_{\tilde{x}^n}^n(s), v, s)/r(\tilde{\Xi}_{\tilde{x}^n}^n(s), v)) ds \right\} \\ &\leq (1 + \gamma) \cdot [\tilde{q}_{\tilde{x}^n}^n(z^n, \sigma, \tau - \sigma; z^n, \beta, \varepsilon) + \eta] + M\gamma\tau. \end{aligned}$$

We now choose the parameter α to be greater than each of the right-hand sides of formulas (6.32), (6.33), and (6.34); namely,

$$\alpha > \delta [BK + 2/3] + 2\sigma\|\beta\| + \varepsilon.$$

Then with probability 1, for any starting point $y \in \mathcal{S}^n$ satisfying $\|y - z\| \leq \delta/3$, this composite control u^n keeps the controlled process $\Xi^n(s)$ a distance less than α from the path $\varphi(s) = z + s\beta$ for s in the entire time interval $[0, \tau]$. It follows that with probability 1 the exit cost $g(\pi(0, \tau, \Xi^n; z, \beta, \alpha))$ equals 0.

The expected integrated running cost over each of the intervals $[\tau - \sigma, \tau]$ and $[0, \sigma]$ is no greater than 2η . When $y \in \mathcal{S}^n$ satisfies $\|y - z\| \leq \delta/3$, equation (6.35) and the fact that with probability 1 the exit cost $g(\pi(0, \tau, \Xi^n; z, \beta, \alpha))$ equals 0 imply that

(6.36)

$$\begin{aligned} & -\frac{1}{n} \log P_{y, 0}^n \left\{ \sup_{s \in [0, \tau]} \|X^n(s) - z - s\beta\| < \alpha \right\} \\ &= q^n(y, 0, \tau; z, \beta, \alpha) \\ &= V^n(y, 0, \tau; z, \beta, \alpha) \\ &\leq E_{z^n, \sigma}^n \left\{ \int_{\sigma}^{\tau - \sigma} \sum_{v \in \mathbb{Z}^d} r(\Xi^n(s), v) \ell(u^n(\Xi^n(s), v, s)/r(\Xi^n(s), v)) ds \right. \\ &\qquad \qquad \qquad \left. + g(\pi(0, \tau, \Xi^n; z, \beta, \alpha)) \right\} + 2\eta \\ &\leq (1 + \gamma) \cdot [\tilde{q}_{\tilde{x}^n}^n(z^n, \sigma, \tau - \sigma; z^n, \beta, \varepsilon) + \eta] + M\gamma\tau + 2\eta. \end{aligned}$$

Since $\lim_{n \rightarrow \infty} z^n = \varphi(t^*)$, we have for any parameter $\kappa \in (0, \varepsilon)$ and all sufficiently large $n \in \mathbb{N}$ the implication

$$\begin{aligned} & \sup_{s \in [\sigma, \tau - \sigma]} \|\tilde{\Xi}_{\tilde{x}^n}^n(s) - \varphi(t^*) - (s - \sigma)\beta\| < \kappa \\ &\Rightarrow \sup_{s \in [\sigma, \tau - \sigma]} \|\tilde{\Xi}_{\tilde{x}^n}^n(s) - z^n - (s - \sigma)\beta\| < \varepsilon. \end{aligned}$$

Hence for $\kappa \in (0, \varepsilon)$ and all sufficiently large $n \in \mathbb{N}$

$$\tilde{q}_{\tilde{x}^n}^n(z^n, \sigma, \tau - \sigma; z^n, \beta, \varepsilon) \leq \tilde{q}_{\tilde{x}^n}^n(z^n, \sigma, \tau - \sigma; \varphi(t^*), \beta, \kappa),$$

and so whenever $y \in \mathcal{S}^n$ satisfies $\|y - z\| \leq \delta/3$ we have from formula (6.36)

$$\begin{aligned} & -\frac{1}{n} \log P_{y,0}^n \left\{ \sup_{s \in [0, \tau]} \|X^n(s) - z - s\beta\| < \alpha \right\} \\ & \leq (1 + \gamma) \cdot [\tilde{q}_{\tilde{x}}^n(z^n, \sigma, \tau - \sigma; z^n, \beta, \varepsilon) + \eta] + M\gamma\tau + 2\eta \\ & \leq (1 + \gamma) \cdot [\tilde{q}_{\tilde{x}}^n(z^n, \sigma, \tau - \sigma; \varphi(t^*), \beta, \kappa) + \eta] + M\gamma\tau + 2\eta \\ & = (1 + \gamma) \cdot \tilde{q}_{\tilde{x}}^n(z^n, \sigma, \tau - \sigma; \varphi(t^*), \beta, \kappa) + (1 + \gamma)\eta + M\gamma\tau + 2\eta. \end{aligned}$$

Since $\varphi(t^*) = z + t^*\beta$, formula (6.11) implies that the points $\varphi(t^*)$ and β both lie in the set $\Gamma(\tilde{x})$. Thus, using the inequality $\|z^n - \varphi(t^*)\| \leq \delta/3$ for all $n \geq N_1$, we can apply the second limit in equation (4.7) to conclude that

$$\begin{aligned} & \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\{y \in \mathcal{S}^n : \|y - z\| \leq \delta/3\}} \left(-\frac{1}{n} \log P_{y,0}^n \left\{ \sup_{s \in [0, \tau]} \|X^n(s) - z - s\beta\| < \alpha \right\} \right) \\ & \leq (1 + \gamma) \cdot \lim_{\kappa \rightarrow 0} \limsup_{n \rightarrow \infty} \tilde{q}_{\tilde{x}}^n(z^n, \sigma, \tau - \sigma; \varphi(t^*), \beta, \kappa) \\ & \quad + (1 + \gamma)\eta + M\gamma\tau + 2\eta \\ & \leq (1 + \gamma) \cdot \lim_{\kappa \rightarrow 0} \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\{w \in \mathcal{S}_x^n : \|w - \varphi(t^*)\| \leq \delta\}} \tilde{q}_{\tilde{x}}^n(w, \sigma, \tau - \sigma; \varphi(t^*), \beta, \kappa) \\ & \quad + (1 + \gamma)\eta + M\gamma\tau + 2\eta \\ & = (1 + \gamma) \cdot \lim_{\kappa \rightarrow 0} \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\{w \in \mathcal{S}_x^n : \|w - \varphi(t^*)\| \leq \delta\}} \\ & \quad \left(-\frac{1}{n} \log P_{\tilde{x}; w, \sigma}^n \left\{ \sup_{s \in [\sigma, \tau - \sigma]} \|\tilde{X}_{\tilde{x}}^n(s) - \varphi(t^*) - (s - \sigma)\beta\| < \kappa \right\} \right) \\ & \quad + (1 + \gamma)\eta + M\gamma\tau + 2\eta \\ & \leq (1 + \gamma) \cdot (\tau - 2\sigma) \cdot L(\tilde{x}, \beta) + (1 + \gamma)\eta + M\gamma\tau + 2\eta \\ & = (1 + \gamma) \cdot (\tau - 2\eta) \cdot L(\tilde{x}, \beta) + (1 + \gamma)\eta + M\gamma\tau + 2\eta. \end{aligned}$$

The upper bound (6.10) now follows by sending $\eta \rightarrow 0$ and then $\alpha \rightarrow 0$.

This completes the proof of the local large deviation principle stated in Proposition 5.1. With this we have finished the proof of the large deviation principle stated in Theorem 4.3.

7. APPENDIX: PROOF OF THE REPRESENTATION FORMULA IN THEOREM 3.3

We take $n \in \mathbb{N}$, $y \in \mathcal{S}^n$, $t \in [0, 1]$, $\beta \in \mathbb{R}^d$, and $\varepsilon \in (0, 1)$ and recall the set \mathcal{B}^n and the numbers $\tilde{t}_1(y)$ and $\tilde{t}_2(y)$, which were defined in the paragraph before equation (3.11) In order to prove Theorem 3.3, we will show the following three statements:

- (1) $V^n(y, t; \beta, \varepsilon) = q^n(y, t; \beta, \varepsilon) = \infty$ if $y \in \mathcal{S}^n \setminus \mathcal{B}^n$ and $t \in [0, 1]$,
- (2) $V^n(y, t; \beta, \varepsilon) = q^n(y, t; \beta, \varepsilon) = \infty$ if $y \in \mathcal{B}^n$ and $t \in [0, \tilde{t}_1(y)] \cup [\tilde{t}_2(y), 1]$,
- (3) $V^n(y, t; \beta, \varepsilon) = q^n(y, t; \beta, \varepsilon) < \infty$ if $y \in \mathcal{B}^n$ and $t \in (\tilde{t}_1(y), \tilde{t}_2(y)) \cap [0, 1]$.

Unless otherwise noted, the phrase “with probability one” in the proof will mean with probability one with respect to $\tilde{E}_{y,t}^n$. In order to simplify the notation,

we will not explicitly indicate the dependence of the various quantities on β and ε , which are fixed throughout this section. Thus we will write $\pi(t, \varphi)$, $V^n(y, t)$, $p^n(y, t)$, and $q^n(y, t)$ instead of $\pi(t, \varphi; \beta, \varepsilon)$, $V^n(y, t; \beta, \varepsilon)$, $p^n(y, t; \beta, \varepsilon)$, and $q^n(y, t; \beta, \varepsilon)$, respectively.

Proof of statement (1). If $y \in \mathcal{S}^n \setminus \mathcal{B}^n$, then there are two possibilities. If $\|y - t\beta\| \geq \varepsilon$ then for any admissible control we have $\pi(t, \Xi^n) \leq 1$ with probability 1. Hence in this case $V^n(y, t) = \infty = q^n(y, t)$. On the other hand, there may be points $y \in \mathcal{S}^n \setminus \mathcal{B}^n$ and $t \in [0, 1]$ such that $\|y - t\beta\| < \varepsilon$. Fix any such y and t . The definition of the set \mathcal{B}^n implies that $p^n(y, t) = 0$. We claim that this implies $V^n(y, t) = \infty = q^n(y, t)$. Let u^n be any admissible control and let Ξ^n be the corresponding controlled process. Assume that $V^n(y, t) < \infty$. Then with probability 1 $\pi(t, \Xi^n) = \infty$ and

$$V^n(y, t) = E_{y,t}^n \left\{ \int_t^1 \sum_{v \in \mathbb{Z}^d} r(\Xi^n(s), v) \ell(u^n(\Xi^n(s), v, s)/r(\Xi^n(s), v)) ds \right\} < \infty.$$

Together with the superlinear growth of the function ℓ and the positive lower bound (4.1) on $r(x, v)$ for all $x \in \mathcal{S}^n$ and $v \in \mathbb{Z}^d$, the last display implies that the distribution on path space induced by $\{\Xi^n(s), s \in [t, 1]\}$ under $E_{y,t}^n$ is absolutely continuous with respect to the distribution induced by $\{X^n(s), s \in [t, 1]\}$ under $E_{y,t}^n$. Because $p^n(y, t) = 0$, we have $\pi(t, X^n) \leq 1$ with probability 1 under $E_{y,t}^n$. This implies that $\pi(t, \Xi^n) \leq 1$ with probability 1, which contradicts $V^n(y, t) < \infty$. We conclude that if $y \in \mathcal{S}^n \setminus \mathcal{B}^n$ and $t \in [0, 1]$, then $V^n(y, t) = \infty = q^n(y, t)$ in both cases. This completes the proof of statement (1).

Proof of statement (2). Since $y \in \mathcal{B}^n$ and $t \in [0, \tilde{t}_1(y)]$ imply $\|y - t\beta\| \geq \varepsilon$, the proof that $V^n(y, t) = \infty = q^n(y, t)$ in this case is the same as the proof in statement (1) of the case where $y \in \mathcal{S}^n \setminus \mathcal{B}^n$, $t \in [0, 1]$, and $\|y - t\beta\| \geq \varepsilon$. Since $y \in \mathcal{B}^n$ and $t \in [\tilde{t}_2(y), 1]$ imply $p^n(y, t) = 0$, the proof that $V^n(y, t) = \infty = q^n(y, t)$ in this case is the same as the proof in statement (1) of the case where $y \in \mathcal{S}^n \setminus \mathcal{B}^n$, $t \in [0, 1]$, and $\|y - t\beta\| < \varepsilon$. This completes the proof of statement (2).

Proof of statement (3). Fix $y \in \mathcal{B}^n$ and $t \in (\tilde{t}_1(y), \tilde{t}_2(y)) \cap [0, 1]$. We first prove that $V^n(y, t) \geq q^n(y, t)$. Let u^n be any admissible control for $V^n(y, t)$ having finite cost and let $\{\Xi^n(s), s \in [t, 1]\}$ denote any corresponding controlled process satisfying $\Xi^n(t) = y$. For any bounded real-valued function f on \mathcal{S}^n the generators of the controlled process are given by the formula

$$(\mathcal{L}^n f)(y, s) \doteq n \sum_{v \in \mathbb{Z}^d} u^n(y, v, s) [f(y + v/n) - f(y)].$$

Since the cost is finite, the definition of $V^n(y, t)$ implies that with probability 1 $g(\pi(t, \Xi^n)) = 0$ and thus $\pi(t, \Xi^n) = \infty$.

For any $r \in (0, \infty)$ and $u \in \mathbb{R}$ the Legendre-Fenchel transform

$$\sup_{\theta \in \mathbb{R}} \{\theta u - r(e^\theta - 1)\} = r \ell(u/r)$$

and convex duality yield for any $r \in (0, \infty)$ and $\theta \in \mathbb{R}$ the relation

$$(7.1) \quad r(e^{-\theta} - 1) = \sup_{u \in \mathbb{R}} \{-u\theta - r \ell(u/r)\}.$$

This in turn yields the inequality

$$(7.2) \quad r(e^{-\theta} - 1) + u\theta + r \ell(u/r) \geq 0$$

for any $r \in (0, \infty)$, $\theta \in \mathbb{R}$, and $u \in \mathbb{R}$.

For $B \in (0, \infty)$ let $q_B^n(y, t)$ satisfy the system of ordinary differential equations

$$(7.3) \quad \frac{\partial q_B^n(y, t)}{\partial t} = \sum_{v \in \mathbb{Z}^d} r(y, v) (\exp(-n [q_B^n(y + v/n, t) - q_B^n(y, t)]) - 1)$$

for $(y, t) \in \mathcal{S}^n \times (0, 1)$ satisfying $\|y - t\beta\| < \varepsilon$; the boundary condition

$$q_B^n(y, t) = B \text{ for } (y, t) \in \mathcal{S}^n \times [0, 1] \text{ satisfying } \|y - t\beta\| \geq \varepsilon;$$

and the terminal condition

$$(7.4) \quad q_B^n(y, 1) = 0 \text{ for } y \in \mathcal{S}^n \text{ satisfying } \|y - \beta\| < \varepsilon.$$

As in Section 3, the boundary condition is valid only in “forward” t -direction. Since the functions $p_B^n(y, t) \doteq -\frac{1}{n} \log q_B^n(y, t)$ satisfy $p_B^n(y, t) \downarrow p^n(y, t)$ for all $y \in \mathcal{S}^n$ and $t \in [0, 1]$, we see that $q_B^n(y, t) \uparrow q^n(y, t)$ for all $y \in \mathcal{S}^n$ and $t \in [0, 1]$.

Recall that $\pi(t, \Xi^n) = \infty$ implies $\|\Xi^n(s) - s\beta\| < \varepsilon$ for all $s \in [t, 1]$. Hence equations (7.2) and (7.3) imply that

$$(7.5) \quad \frac{\partial q_B^n(\Xi^n(s), s)}{\partial s} + \mathcal{L}^n q_B^n(\Xi^n(s), s) + \sum_{v \in \mathbb{Z}^d} r(\Xi^n(s), v) \ell(u^n(\Xi^n(s), v, s)/r(\Xi^n(s), v)) \geq 0$$

for all $s \in [t, 1)$ and with probability 1. Integrating s from t to 1 and taking the expected value gives

$$(7.6) \quad E_{y,t}^n \left\{ \int_t^1 \left(\frac{\partial q_B^n(\Xi^n(s), s)}{\partial s} + \mathcal{L}^n q_B^n(\Xi^n(s), s) \right) ds \right\} + E_{y,t}^n \left\{ \int_t^1 \sum_{v \in \mathbb{Z}^d} r(\Xi^n(s), v) \ell(u^n(\Xi^n(s), v, s)/r(\Xi^n(s), v)) ds \right\} \geq 0.$$

Using the martingale property that was assumed as part of the definition of admissibility and the terminal condition (7.4), we obtain

$$E_{y,t}^n \left\{ \int_t^1 \left(\frac{\partial q_B^n(\Xi^n(s), s)}{\partial s} + \mathcal{L}^n q_B^n(\Xi^n(s), s) \right) ds \right\} = -q_B^n(y, t).$$

Therefore equation (7.6) implies that

$$\bar{E}_{y,t}^n \left\{ \int_t^1 \sum_{v \in \mathbb{Z}^d} r(\Xi^n(s), v) \ell(u^n(\Xi^n(s), v, s)/r(\Xi^n(s), v)) ds \right\} \geq q_B^n(y, t).$$

Sending $B \rightarrow \infty$, we obtain

$$\bar{E}_{y,t} \left\{ \int_t^1 \sum_{v \in \mathbb{Z}^d} r(\Xi^n(s), v) \ell(u^n(\Xi^n(s), v, s)/r(\Xi^n(s), v)) ds \right\} \geq q^n(y, t).$$

Hence the definition of $V^n(y, t)$ as a minimal cost function allows us to conclude that

$$V^n(y, t) \geq q^n(y, t).$$

In order to prove that $V^n(y, t) \leq q^n(y, t)$, we first note that for any $\theta \in \mathbb{R}$ the supremum in formula (7.1) is attained at the unique value $u \doteq r e^{-\theta}$. This suggests that we can define an optimal control by letting

$$(7.7) \quad u^n(y, v, s) \doteq r(y, v) \exp(-n[q^n(y + v/n, s) - q^n(y, s)])$$

if $y \in \mathcal{B}^n$ and $s \in (\tilde{t}_1(y), \tilde{t}_2(y)) \cap [0, 1]$ and by letting $u^n(y, v, s) = 0$ for all other values of $y \in \mathcal{S}^n$ and $t \in [0, 1]$. We define $\exp(-\infty) \doteq 0$. Since $q^n(y, t) < \infty$ for $s \in (\tilde{t}_1(y), \tilde{t}_2(y)) \cap [0, 1]$, $u^n(y, v, s)$ is well defined and measurable.

For $\eta \in (0, 1)$ we define the stopping time τ_η by the formula

$$(7.8) \quad \tau_\eta \doteq \inf\{s \geq t : s \geq \tilde{t}_2(\Xi^n(s)) - \eta \text{ or } \tilde{t}_1(\Xi^n(s)) \geq s\} \wedge 1.$$

For all $s \in [t, \tau_\eta]$ and all ω in the underlying probability space, we have

$$\tilde{t}_1(\Xi^n(s)) < s < \tilde{t}_2(\Xi^n(s)) - \eta.$$

This fact and the definition (7.7) imply that the controls $u^n(\Xi^n(s), v, s)$ are uniformly bounded and continuous for all $s \in [t, \tau_\eta]$. This uniquely defines the distribution of $\{\Xi^n(s), s \in [t, \tau_\eta]\}$.

For $y \in \mathcal{B}^n$ and $t \in (\tilde{t}_1(y), \tilde{t}_2(y)) \cap (0, 1)$ the function $q^n(y, t; \beta, \varepsilon)$ satisfies the system of ordinary differential equations given in equation (3.11). Hence for all $s \in [t, \tau_\eta]$ and all ω in the underlying probability space, our choice of control implies that

$$\begin{aligned} \frac{\partial q^n(\Xi^n(s), s)}{\partial s} + \mathcal{L}^n q^n(\Xi^n(s), s) \\ + \sum_{v \in \mathbb{Z}^d} r(\Xi^n(s), v) \ell(u^n(\Xi^n(s), v, s)/r(\Xi^n(s), v)) = 0. \end{aligned}$$

Integrating s from t to τ_η and taking the expected value gives

$$(7.9) \quad \begin{aligned} \bar{E}_{y,t}^n \left\{ \int_t^{\tau_\eta} \left(\frac{\partial q^n(\Xi^n(s), s)}{\partial s} + \mathcal{L}^n q^n(\Xi^n(s), s) \right) ds \right\} \\ + \bar{E}_{y,t} \left\{ \int_t^{\tau_\eta} \sum_{v \in \mathbb{Z}^d} r(\Xi^n(s), v) \ell(u^n(\Xi^n(s), v, s)/r(\Xi^n(s), v)) ds \right\} = 0. \end{aligned}$$

Since the jump rates $u^n(\Xi^n(s), v, s)$ are uniformly bounded for all $s \in [t, \tau_\eta]$ and all ω in the underlying probability space, the quantity

$$\int_t^{\tau_\eta \wedge \alpha} \left(\frac{dq^n(\Xi^n(s), s)}{ds} - \frac{\partial q^n(\Xi^n(s), s)}{\partial s} - \mathcal{L}^n q^n(\Xi^n(s), s) \right) ds$$

is a martingale in α for $\alpha \in [t, 1]$. Thus equation (7.9) implies that

$$(7.10) \quad E_{y,t}^n \left\{ \int_t^{\tau_\eta} \sum_{v \in \mathbb{Z}^d} r(\Xi^n(s), v) \ell(u^n(\Xi^n(s), v, s)/r(\Xi^n(s), v)) ds \right\} = q^n(y, t) - E_{y,t}^n \{q^n(\Xi^n(\tau_\eta), \tau_\eta)\}.$$

Recall that in the statement (3) now under consideration $q^n(y, t) \in [0, \infty)$ and $q^n(y, s) \geq 0$ for all $y \in \mathcal{S}^n$ and $s \in [0, 1]$. Hence the last equation implies that

$$(7.11) \quad \sup_{\eta \in (0,1)} E_{y,t}^n \left\{ \int_t^{\tau_\eta} \sum_{v \in \mathbb{Z}^d} r(\Xi^n(s), v) \ell(u^n(\Xi^n(s), v, s)/r(\Xi^n(s), v)) ds \right\} < \infty$$

and that

$$(7.12) \quad \sup_{\eta \in (0,1)} E_{y,t}^n \{q^n(\Xi^n(\tau_\eta), \tau_\eta)\} < \infty.$$

Define $\tau_0 \doteq \lim_{\eta \downarrow 0} \tau_\eta$. Equation (7.11) and the monotone convergence theorem imply that

$$E_{y,t}^n \left\{ \int_t^{\tau_0} \sum_{v \in \mathbb{Z}^d} r(\Xi^n(s), v) \ell(u^n(\Xi^n(s), v, s)/r(\Xi^n(s), v)) ds \right\} < \infty.$$

The superlinear growth of the function ℓ and the positive lower bound (4.1) on $r(x, v)$ for $x \in \mathcal{S}^n$ and $v \in \Sigma(y)$ then yield that

$$E_{y,t}^n \left\{ \int_t^{\tau_0} \sum_{v \in \mathbb{Z}^d} u^n(\Xi^n(s), v, s) ds \right\} < \infty.$$

This in turn implies that with probability 1 $\Xi^n(s)$ jumps only a finite number of times for s in the interval $[t, \tau_0)$, Hence the quantity $Z \doteq \lim_{\eta \downarrow 0} \Xi^n(\tau_\eta)$ is well-defined with probability 1. By Fatou's Lemma and equation (7.12), we obtain

$$E_{y,t}^n \{q^n(Z, \tau_0)\} < \infty.$$

Now the definition of the stopping times τ_η implies either that $\tau_0 = 1$ and $\|Z - \beta\| < \varepsilon$, that $\tau_0 = \tilde{i}_2(Z)$, or that $\tau_0 = \tilde{i}_1(Z)$. However, since $q^n(y, s) = \infty$ if $s \in \{\tilde{i}_1(y), \tilde{i}_2(y)\}$, we conclude that $\tau_0 = 1$ and $\|Z - \beta\| < \varepsilon$ with probability 1 and therefore that $E_{y,t}^n \{q^n(Z, \tau_0)\} = 0$. Since $\tau_0 = 1$, one can check that Ξ^n satisfies the martingale property and the uniqueness in distribution property required for u^n to be an admissible control and furthermore that

$\pi(t, \Xi^n) = \infty$ with probability 1. We conclude from equation (7.10) that

$$\bar{E}_{y,t}^n \left\{ \int_t^1 \sum_{v \in \mathbb{Z}^d} r(\Xi^n(s), v) \ell(u^n(\Xi^n(s), v, s)/r(\Xi^n(s), v)) ds + g(\pi(t, \Xi^n)) \right\} \\ = q^n(y, t),$$

The definition of $V^n(y, t)$ as a minimal cost function then implies that

$$V^n(y, t) \leq q^n(y, t).$$

This completes the proof of Theorem 3.3. \square

ACKNOWLEDGEMENT

We thank a referee for a careful reading of this paper and a number of helpful suggestions.

REFERENCES

1. R. Anderson and S. Orey, *Small random perturbations of dynamical systems with reflecting boundary*, Nagoya Math. J. **60** (1976), 189–216.
2. P. Billingsley, *Convergence of probability measures*, Wiley, New York, 1968.
3. J.-D. Deuschel and D. W. Stroock, *Large deviations*, Academic Press, San Diego, 1989.
4. P. Dupuis, *Large deviations analysis of reflected diffusions and constrained stochastic approximation algorithms in convex sets*, Stochastics **21** (1987), 63–96.
5. P. Dupuis and R. S. Ellis, *Large deviations for Markov processes with discontinuous statistics, II: Random walks*, Probab. Theory Related Fields **91** (1992), 153–194.
6. ———, *A weak convergence approach to the theory of large deviations*, Wiley, New York, 1995 (to appear).
7. P. Dupuis, R. S. Ellis, and A. Weiss, *Large deviations for Markov processes with discontinuous statistics, I: General upper bounds*, Ann. Probab. **19** (1991), 1280–1297.
8. P. Dupuis and H. Ishii, *On Lipschitz continuity of the solution mapping to the Skorokhod problem, with applications*, Stochastics Stochastic Rep. **35** (1991), 31–62.
9. P. Dupuis, H. Ishii, and H. M. Soner, *A viscosity solution approach to the asymptotic analysis of queueing systems*, Ann. Probab. **18** (1990), 226–255.
10. G. Kieffer, *The large deviation principle for two-dimensional stable systems*, Ph. D. Thesis, Univ. of Massachusetts, 1995.
11. A. A. Mogulskii, *Large deviations for trajectories of multi-dimensional random walks*, Theory Probab. Appl. **21** 300–315 (1976).
12. Y. W. Park, *Large deviation theory for queueing systems*. Ph.D. thesis, Virginia Polytechnic Institute and State University, 1991.
13. D. W. Stroock, *An introduction to the theory of large deviations*, Springer-Verlag, New York, 1984.
14. P. Tsoucas, *Rare events in series of queues*, J. Appl. Probab. **29** (1992) 168–175.

LEFSCHETZ CENTER FOR DYNAMICAL SYSTEMS, DIVISION OF APPLIED MATHEMATICS, BROWN UNIVERSITY, PROVIDENCE, RI 02912

E-mail address: dupuis@cfm.brown.edu

DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF MASSACHUSETTS, AMHERST, MASSACHUSETTS 01003

E-mail address: rsellis@math.umass.edu