CCIT Report #375
March 2002

Large deviations with diminishing rates*

Adam Shwartz
Electrical Engineering
Technion—Israel Institute of Technology
Haifa 32000, Israel

Alan Weiss
Bell Laboratories, Lucent Technologies
Murray Hill, NJ

March 2002

Abstract

The theory of large deviations for jump Markov processes has been generally proved only when jump rates are bounded below, away from zero [2, 5, 7]. Yet various applications of interest do not satisfy this condition. We describe several classes of models where jump rates diminish to zero in a Lipschitz continuous way. Under appropriate conditions, we prove that the sample path large deviations principle continues to hold. Under our conditions, the rate function remains an integral over a local rate function, which retains its standard representation.

1 Introduction

One of the common technical assumptions in existing large deviations theory for jump Markov processes is that jump rates are bounded below, away from zero [2, 5, 7]. This is not merely a technical assumption: if the rates may go down to zero, the process may get stuck at a point, or it may or may not be possible to reach certain regions. We illustrate these issues below through

*Research supported in part by the United States—Israel Binational Science Foundation grant no. BSF 1999179. Research of the first author supported also in part by the fund for promotion of research at the Technion.

examples. From a technical point of view, this condition is required in order to obtain smoothness of the local rate function.

However, in many applications this condition is violated. In this paper we prove the large deviations principle for the sample paths of a wide class of such models. Under our conditions, the usual Markovian integral representation of the rate function continues to hold, with the same variational formula for the integrand, as in the case with rates bounded away from zero. We allow the boundary defined by the region where some rates go to zero to be fairly general; in contrast with existing large deviations theory which deals only with "flat" boundaries, our boundaries are quite general.

Applications for which rates are not bounded away from zero fall roughly into two categories. In the first, the rates are proportional to the occupancy of the system, and thus go to zero when the system empties. In the second, a control is designed in order to avoid, say, overflows. "Soft controls" are characterized by continuous jump rates, and these controls may well have rates become zero. Another example of a soft control is when additional processing power may be gradually added to a system when it grows beyond a certain threshold. Here are some specific examples of these applications.

Example A. In an M/M/∞ queue, the service rate is linear in the queue size, so that as the queue empties, the rate goes to zero. In this case it is not possible to continue reducing the queue size: the rate diminishes as we approach a boundary of the state space. This type of behavior is one of our main motivations, as there are many multi-dimensional models exhibiting this behavior [6].

Example B. Consider an $M/M/1$ queue scaled by n: arrivals occur at rate $n\lambda$, services at rate $n\mu$, and the queue size $x(t)$ is scaled to $z_n(t) = x(t)/n$. Suppose that there is an auxiliary $M/M/\infty$ server, also with parameter μ, that kicks in when the scaled queue size $z_n(t) > 1$, and that each service of this auxiliary server kicks out a pair of customers. The rate of departure of pairs is thus $\mu \cdot (z_n(t) - 1)$ whenever $z_n(t) > 1$, and 0 otherwise. Since the rate of jumps of size (-2) diminishes to zero at $z_n = 1$, this case falls outside the current theory. One key feature of this model is that the set of possible jump directions (the positive cone spanned by the jump directions) does not change as the jump rate diminishes to zero.

Example C. Consider a two queue system with "soft priority" to queue 1. We suppose the two queues behave like independent M/M/1 queues, with arrival rates 0.5 and service rate 1, as long as the (scaled) size of queue 1 is below 1. When the first queue is larger than one, the server of queue 2 enters a processor sharing mode, so that the service rate to queue 1 increases linearly with queue size until its queue size is 2, and the rate remains at 2 when the first queue is larger, while service rate at queue 2 decreases linearly with the size of queue 1 until no service is offered at all. In this case the service rate at queue 2 goes down and is zero on $\{\vec{x} : x_1 \geq 2\}$. In contrast with the Example A, this region can be entered with positive probability, so that it is not delineated by a boundary. In contrast with our other examples, we do *not* develop a theory that covers this case, although the methods we develop suffice for the analysis of simple models of this type.

Example D. We describe a simplified model of connection admission control (CAC), based on an idea of David Tse [8]. Suppose that customers arrive at an infinite-server queue with Poisson rate $n\lambda$. Arrivals may be turned away (blocked), according to rules given in the next paragraph. Accepted customers depart the queue at rate μ. We let $q(t)$ represent the number in the system at time t. Customers who are accepted (not blocked) have two states, represented by 0 and 1. Accepted customer i moves from $x_i = 0$ to $x_i = 1$ with rate α, and from $x_i = 1$ to $x_i = 0$ with rate β. The bandwidth customer i uses at time t is $x_i(t)$; the total bandwidth in use at time t is thus $b(t) \triangleq \sum_{i=1}^{q(t)} x_i(t)$.

Now for the problem. Suppose that the capacity of the system is nC, so that there is trouble if $\sum_i x_i(t) > nC$. We try to ensure that this happens with small probability by accepting connections only when the scaled system state $\vec{z}_n(t) \triangleq 1/n(b(t), q(t))$ is in a fixed region G. For some fixed $\delta > 0$ we define the probability that a connection is accepted by

$$\mathbb{P}(\text{accept}) = \begin{cases} 1 & \text{if } \vec{z}_n(t) \in G \text{ and } d[\vec{z}_n(t), \partial G] > \delta \\ \frac{1}{\delta}\text{dist}[\vec{z}_n(t), \partial G] & \text{if } \vec{z}_n(t) \in G \text{ and } d[\vec{z}_n(t), \partial G] \leq \delta \\ 0 & \text{otherwise} \end{cases}$$

We may try to design the set G so that the cheapest rate function for a path from $\vec{z}_n(0)$ to the set $b(t)/n = C$ is the same for all starting points $\vec{z}_n(0) \in \partial G$. In conclusion, we have described a Markov model of CAC that fits into the theory we develop in this paper, although we do not carry out the analysis and design of the appropriate region G for this model.

Example E. This example is to show some potential pitfalls in the theory of diminishing rates. Consider a pure birth process with $\lambda(x) = |x|$. This process moves to the right and, since $\lambda(0) = 0$, if $x(0) < 0$ is an integer, then $x(t) \leq 0$ for all t. Using the formulas for the local rate functions that hold true when the rates are bounded below, a formal calculation (detailed below in Section 7) of the rate function $I(r)$ for the path $r(t) = t - 0.5$, $t \in [0, 1]$ yields $I(r) < \infty$, implying that $\mathbb{P}\left(\text{process} \approx r(t)\right) \approx e^{-nI(r)} > 0$. It is clear that the probability of the process following near this path is exactly 0, so the formal calculation of the rate is incorrect. However, if the process $x(t)$ starts with a non-integer value, then the probability is strictly positive, and it turns out that in this case the formal calculation is correct, although we will not detail that straightforward calculation in this paper. In more generality, if the rates $\lambda(x)$ are bounded below, then the sequence of processes with rates $\lambda^n(x) = n\lambda(x)+1$ is exponentially equivalent to the case with $\lambda^n(x) = n\lambda(x)$ so that, on the large deviations scale, their behavior is identical. However, for the decreasing rate case described above it leads to completely different behavior, since now it becomes possible to cross the $x = 0$ barrier. The upshot is that the form of the rate function we derive is in some sense less robust than it is under the usual condition that the rates are bounded away from 0.

We note that our analysis holds for processes confined to a convex set by having jump rates, in directions heading out of the set, diminish to zero at the boundary, as in Example D. Thus we establish, for the first time, a sample-path large deviations principle in the case of curved boundaries. Previous published work [7, Chapter 8], [3, 2, 5] dealt only with flat boundaries; there is some unpublished work dealing with curved boundaries with some restrictions. Our general approach is based on [7]; indeed, we use this as the source of many of our lemmas, and sometimes prove theorems by giving only the changes necessary to use arguments in [7].

In general, there is no exponential equivalence between processes with log-bounded rates and those whose rates may vanish, as Example E clearly shows. In fact, it is natural to approximate a process with rates that go to zero by imposing a small lower bound, say ε, on the rates, and there is an obvious coupling between these two processes. However, it is not hard to see that this coupling does not provide an exponential approximation. Therefore, we resort to the more technical approach of following the steps of [7].

The paper is organized as follows. In the next Section 2 we set up the notation and describe the problem as well as our main results: Theorem 2 for processes with boundaries, where some jump rates go to zero at the boundary, and Theorem 3 for processes with diminishing rates but where, at each point, the jump directions associated with positive rates span $\mathbb{R}^d$. Corollary 4, § 2 shows that these two results can be combined, to cover a large class of models exhibiting small rates both near boundaries and in the interior. Finally, in Corollary 8 we remove the technical assumption that the set of interest is compact, to obtain our results under the weakest assumptions. This last extension is a consequence of Corollary 7, a result of independent interest, on the exponential tightness of models with rates that grow at most linearly.

In Section 3 we develop some preliminary results. Next in Section 4 we prove the large deviations upper bound for the case that rates diminish at a boundary. In Section 5 we establish the corresponding lower bound. In

Section 6 we state and prove the large deviations principle when some rates become zero in the interior of the region, for models such as Example B of the introduction.

In Section 7 we give a simple sufficient condition for when the rate function is finite for paths that approach a boundary; this shows that many boundaries may be reached with probability that decreases to zero at an exponential rate, as opposed to a superexponential rate. This condition is also virtually necessary in the one dimensional case. Section 8 summarizes our results, and sketches open problems. The appendix contains a technical lemma that allows us to extend Lipschitz continuous jump rates from regions that are unions of convex sets to all of $\mathbb{R}^d$.

2 Assumptions and main results

We study jump Markov processes with boundaries where some jump rates become zero. The structure of the processes is simple, and is detailed first. The structure of the boundaries of the regions of interest is more complicated, and is deferred for a few paragraphs. Using the notation of [7], our model of jump processes is specified through k jump directions $\{\vec{e}_j\}_{j=1}^k$ and their respective Poisson jump rates $\{\lambda_j(\vec{x})\}$, which are defined for all $\vec{x} \in G$. We assume that the jump rates are Lipschitz continuous functions and, without loss of generality, that the positive cone (defined in (2.3)) spanned by the $\{\vec{e}_j\}$ is $\mathbb{R}^d$. We call such a process $x(t)$. Previous studies have assumed that the the jump rates are uniformly bounded away from 0; the sole novelty of this paper is the relaxation of that condition.

We scale the process $x(t)$ in space and time to $z^n(t)$ as follows. All jump rates are multiplied by a scaling parameter n, and all jump sizes are divided by n; in other words, for $z^n(t)$, jump $\vec{e}_j$ becomes $\vec{e}_j/n$, and occurs at rate $n\lambda_j$. The generators for the original and scaled process are thus given respectively by

$$\mathcal{L}f(\vec{x}) \stackrel{\triangle}{=} \sum_j \lambda_j(\vec{x})(f(\vec{x}+\vec{e}_j) - f(\vec{x})) \tag{2.1}$$

$$\mathcal{L}^n f(\vec{x}) \stackrel{\triangle}{=} \sum_j n\lambda_j(\vec{x}) \left(f\left(\vec{x} + \frac{\vec{e}_j}{n}\right) - f(\vec{x})\right). \tag{2.2}$$

We now begin our description of the boundaries with some notation and definitions. We denote by x_j the jth coordinate of a vector $\vec{x}$, by ∂S the boundary of a set S, by S^o the interior of S, by $B(x,\delta)$ the ball of radius δ centered at x, and by $d(\vec{x},\vec{y})$ the Euclidean distance between $\vec{x}$ and $\vec{y}$. Given a set of vectors $\{\vec{u}_j\}$, the positive cone spanned by the vectors is

$$\mathcal{C}\{\vec{u}_j\} \triangleq \left\{ \vec{v} : \text{there exist } \alpha_j \geq 0 \text{ with } \vec{v} = \sum_j \alpha_j \vec{u}_j \right\}. \tag{2.3}$$

The cone generated by the positive jump rates at $\vec{x}$ is denoted by

$$\mathcal{C}_x \triangleq \mathcal{C}\left\{\vec{e}_j : \lambda_j(\vec{x}) > 0\right\}. \tag{2.4}$$

To motivate our assumptions, suppose G is bounded and convex; then it as well as its boundary are compact. Therefore, they can be covered with a finite number of open balls, and in particular we can achieve this covering with some balls centered on the boundary, and the rest having no intersection with the boundary. Moreover, at each point of ∂G we can fit a cone with its narrow part contained entirely in G. More precisely, we say that G has an interior cone property if there are numbers $\varepsilon > 0$, $\beta > 0$, and for every $\vec{x} \in \partial G$ there is a vector $\vec{v}$ such that for each $t \in (0, \varepsilon)$ we have $B(\vec{x} + t\vec{v}, \beta t) \subset G$.

Assumption 1 *The set G is compact, the closure of its interior. There exist positive $\eta, \epsilon, \gamma, \delta_0, \beta$, vectors $\vec{v}_i$ and open balls B_i so that*
(i) $\{B_i = B(\vec{x}_i, r_i), i = 1, \ldots I_1\}$ covers ∂G and $\vec{x}_i \in \partial G$.
(ii) $\{B_i = B(\vec{x}_i, r_i), i = 1, \ldots I\}$ covers G.
(iii) G satisfies an interior cone condition with parameters ε and β. The vectors $\vec{v}$ for the interior cone may be taken as constant, $\vec{v}_i$, in each region B_i. Moreover, if $d(\vec{x}, \partial G) < \gamma$ then $d(\vec{x} + t\vec{v}_i, \partial G) > \beta t$ and is monotone increasing for $0 \leq t \leq \epsilon$.
(iv) For any $\vec{x} \in G$ we have $B(\vec{x}, \delta_0) \cap G \subset B_i$ for some i.

To illustrate that these assumptions are quite weak, we consider the following class. A set S is called star-like with respect to a point $x \in S$ if for any point $y \in S$ the closed line segment between x and y lies in S.

Lemma 1 *Suppose G is compact and that there exists a ball $B \subset S$ so that S is star like with respect to each $x \in B$. Then Assumption 1 holds.*

In particular, any compact convex set with non-empty interior satisfies the assumption of the lemma.

Proof. The first statement holds since G is compact and contains a ball. Since the boundary is compact, we can cover ∂G with balls B centered on the boundary. Cover the interior with balls contained in G^o. Now extract a finite cover, and (i)–(ii) are established.

Take a ball $B(\vec{x}_0, \delta)$, $\delta \leq 2$, contained in B. Without loss of generality we may assume that $\max_i r_i < \frac{\delta}{2}$. Suppose $\vec{x} \in B_j$. Then the convex hull of $B_j \cup B(\vec{x}_0, \delta)$ is contained in G, and so the first part of (iii) follows. For the second part take $\vec{v}_j = \frac{\vec{x}_0 - \vec{x}_j}{2\|\vec{x}_0 - \vec{x}_j\|}$. The last part of (iii) follows from the first since there is a minimal angle to the (finite number of) cones. Finally, (iv) follows by compactness: assume the contrary and take a sequence of points for which the largest ball is of size 2^{-n}. Then the limit point cannot belong to any B_i, a contradiction. ∎

A second property that holds easily when G is convex is that we can extend the jump rates from G to $\mathbb{R}^d$ as follows. Since G is convex, then for each point $\vec{x} \notin G$ there is a unique point $p(\vec{x}) \in \partial G$ that is closest to G: it is the projection of $\vec{x}$ on G. The definition of $\lambda_j(\vec{x})$ can therefore be extended to $\mathbb{R}^d$ by setting $\lambda_j(\vec{x}) \stackrel{\triangle}{=} \lambda_j(p(\vec{x}))$. In the Appendix we show that if the $\lambda_i(\vec{x})$ are Lipschitz continuous for $\vec{x} \in G$, then they are Lipschitz when extended in this way.

The choice of vectors $\vec{v}_i$ in Assumption 1.(iii) is obviously not unique. Below we make the assumption that this choice can be made so that the vectors are consistent with the directions of increase of the diminishing rates.

The following assumption concerns the case where rates diminish near a boundary.

Assumption 2 *The rates and jump directions satisfy the following.*

A. *There is a constant K_λ such that $|\lambda_j(\vec{x}) - \lambda_j(\vec{y})| \leq K_\lambda |\vec{x} - \vec{y}|$. Moreover, the rates can be extended to a δ neighborhood of G, so that the Lipschitz property continues to hold.*

B. *For each $\vec{x} \in \partial G$ there are $\varepsilon_i > 0$ so that $\vec{y} \in \mathcal{C}_x$ together with $|\vec{y}| < \varepsilon_1$ implies $\vec{x} + \vec{y} \in G$.*

C. *$\lambda_i(\vec{x}) > 0$ for all i and $\vec{x} \in G^o$. Moreover, γ, ϵ and $\vec{v}_i$ of Assumption 1 can be chosen so that*

$$\vec{v}_i \in \mathcal{C}\left\{ \vec{e}_j : \inf_{\vec{x} \in B_i} \lambda_j(\vec{x}) > \gamma \right\}$$

and if $\vec{x} \in B_i$, $d(\vec{x}, \partial G) < \gamma$ and $\lambda_j(\vec{x}) < \gamma$ then $\lambda_j(\vec{x} + \alpha\vec{v})$ is monotone increasing in α for $0 < \alpha < \varepsilon$.

D. *$\mathcal{C}\{\vec{e}_j\} = \mathbb{R}^d$.*

Note that assumptions C and D together show $\vec{x} \in G^o$ implies $\mathcal{C}_x = \mathbb{R}^d$.

When rates diminish in the interior, Assumptions 2.B–C are not relevant, and Assumption 2.D is replaced with

Assumption 3 *$\mathcal{C}_x = \mathbb{R}^d$ for all $\vec{x}$.*

This assumption is used for Theorem 3, which does not require all the parts of Assumption 2.

Note that our assumptions allow the process to jump out of G: however, by Assumption 2, at any given point in G—including the boundary—if the jump size is small enough, then one jump will not cause the process to exit G. In particular, it is not possible to jump in a direction parallel to ∂G in directions where G is strictly convex. Finally note that Assumptions 2.B–C imply that rates for jumps out of G must decrease to 0 as the boundary is approached.

As discussed above, Assumption 1 always holds for convex sets, and the second part of Assumption 2.A follows from the first for convex sets. In the Appendix we discuss these properties for unions of convex sets. It is not hard to show that if there is a δ_0 so that for each $\vec{x} \in \partial G$ there is a closed ball $B(\vec{y}, \delta_0)$ with $B(\vec{y}, \delta_0) \cap G = \vec{x}$ then the second part of Assumption 2.A holds, as we can define a projection locally. Note that in our example of the non-robustness of the rate function (Introduction, Example E), the set G is the union of the convex sets $(-\infty, 0]$ and $[0, \infty)$ which do not satisfy the intersection property of Assumption 1 since for the ball containing 0 there can be no $\vec{v}$ satisfying 1.(iii). Consequently, our theory does not apply.

For $\vec{x}, \vec{y} \in \mathbb{R}^d$ and measurable $\vec{r} : [0,T] \to G$ define

$$\ell(\vec{x}, \vec{y}) \triangleq \sup_{\vec{\theta} \in \mathbb{R}^d} \left(\langle \vec{\theta}, \vec{y} \rangle - \sum_j \lambda_j(\vec{x}) \left(e^{\langle \vec{\theta}, \vec{e}_j \rangle} - 1 \right) \right) \tag{2.5}$$

$$I_{[0,T]}(\vec{r}) \triangleq \begin{cases} \int_0^T \ell(\vec{r}(t), \vec{r}^{\,\prime}(t))\, dt & \text{if } \vec{r}(t) \text{ is absolutely continuous,} \\ \infty & \text{otherwise.} \end{cases} \tag{2.6}$$

We can now state our main results. Let D denote the space of bounded, right continuous functions from $[0,T]$ to $\mathbb{R}^d$, possessing left-hand limits, and let D_s denote the space D endowed with the sup norm topology. For any measurable set S in D_s denote

$$I_{[0,T]}(S) \triangleq \inf\{I_{[0,T]}(\vec{r}) : \vec{r} \in S,\ \vec{r}(0) = \vec{x}\}.$$

Recall that a non negative function $I_{[0,T]}$ is called a good rate function if its level sets $\{\vec{r} : I_{[0,T]}(\vec{r}) < \alpha\}$ are compact (for each α). Our main result concerning rates that diminish towards a boundary is

Theorem 2 *Let Assumptions 1–2 hold and consider the sequence $\{\vec{z}^n\}$ of processes taking values in D_s. This sequence satisfies a large deviations principle with good rate function $I_{[0,T]}$. That is, for each $T > 0$, closed set $C \in D_s$, and open set $O \in D_s$, and each point $\vec{x} \in G$,*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_x\left(\vec{z}^n \in C\right) \leq -I_{[0,T]}(C) \tag{2.7}$$

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}_x\left(\vec{z}^n \in O\right) \geq -I_{[0,T]}(O). \tag{2.8}$$

Suppose now that $G = \mathbb{R}^d$; the process has no boundaries. When the rates diminish in the interior without changing the positive cone of jump direction, we have

Theorem 3 *Let Assumptions 2.A and 3 hold and consider the sequence $\{\vec{z}^n\}$ of processes taking values in D_s. This sequence satisfies a large deviations principle with good rate function $I_{[0,T]}$.*

The two results can be combined as follows. The assumptions of Theorem 2 are imposed near the boundary, while the assumptions of Theorem 3 are imposed away from the boundary. Thus rates may go to zero in the interior, but the cone can only change on the boundary.

Corollary 4 *Let B be an open set so that $\overline{B} \subset G^o$. Let Assumptions 1, 2.A and 2.B hold in G, let Assumption 2.C hold in $G \setminus B$ and Assumption 3 holds for $\vec{x} \in \overline{B}$. Consider the sequence $\{\vec{z}^n\}$ of processes taking values in D_s. This sequence satisfies a large deviations principle with good rate function $I_{[0,T]}$.*

Proof. Use Theorems 2 and 3 in each region separately. Then connect the results exactly as in the proof of Lemma 18: see also the comments at the end of the proof of the lower bound. ∎

Comment: as in [7], we move freely between a jump process and its piecewise linear interpolation, without changing notation. As shown there, the two processes are exponentially equivalent under the sup norm (as their distance at any time is at most one jump). We therefore work with interpolated processes, which are piecewise linear (and in particular Lipschitz continuous) and for which the notion of compactness is easier to handle.

Corollary 5 *Theorems 2, 3 and Corollary 4 hold as stated if we endow the space D with the Skorohod J_1 topology.*

Proof. Consider the linearly interpolated version of the process. The identity map from D_s to D with the Skorohod topology is continuous. Therefore by the contraction principle, the theorem holds with the same rate function. Since the distance between the jump process and the interpolated process is at most $1/n$, the two are exponentially equivalent under the Skorohod J_1 metric, and the result is established. ■

Corollary 7, a result of independent interest, shows that linear growth rates guarantee exponential tightness. This implies that is suffices to prove the upper bound for compact sets, namely to establish the weak large deviations principle. Thus, the compactness assumption in Assumption 1 is not necessary: it suffices that the assumptions hold for bounded subsets of G.

Lemma 6 *Assume that the rates $\lambda_i(\vec{x})$ have linear growth: $\lambda_i(\vec{x}) \leq K(1 + |\vec{x}|)$. Then, uniformly for $\vec{x}$ in compact sets,*

$$\lim_{r\to\infty} \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_x \left(\sup_{0\leq t\leq T} |\vec{z}^n(t)| > r \right) = -\infty \ . \tag{2.9}$$

Proof. Fix b_0 and a point $\vec{x}$ with $|\vec{x}| \leq b_0$. Define a process y^n by setting $y^n(0) = b_0$ and having y^n increase by $K_e = \max_j |\vec{e}_j|$ with every jump of $\vec{z}^n$. Then, since $y^n(t)$ is increasing,

$$\sup_{0\leq t\leq T} |\vec{z}^n(t)| \leq b_0 + \frac{y^n(T)}{n} \ . \tag{2.10}$$

Now couple y^n to a collection of independent Yule processes x_i, which are pure birth processes, with $x_i(0) = 1$ and jump rates $K_1 x_i$ where $K_1 \stackrel{\triangle}{=} K_\lambda k K_e$, so that

$$y^n(t) \leq b_0 K_e \sum_1^n x_i(t) \qquad n \geq b_0 \ . \tag{2.11}$$

Using [7, Corollary 14.14], it follows that there exists a good rate function ℓ_T so that

$$\ell_T \ \left(e^{K_1 T}\right) = 0, \quad \lim_{m\to\infty} \ell_T(am) = \infty \tag{2.12}$$

for all $a > 0$, and such that $\ell_T(a)$ is the rate function for $x_i(t)$. That is, for each $a > e^{K_1 T}$ we have

$$\mathbb{P}\left(\sum_{i=1}^n x_i(t) \geq ap \right) \leq e^{-n\ell_T(a)}. \tag{2.13}$$

Therefore,

$$\limsup_n \frac{1}{n} \log \mathbb{P}_x \left(\sup_{0\le t\le T} |\vec{z}^n(t)| > r \right) \le \limsup_n \frac{1}{n} \log \mathbb{P}_x \left(\sum_1^n x_i(t) > \frac{(r-b_0)n}{K_2} \right) \tag{2.14}$$

$$= -\ell_T \left(\frac{(r-b_0)}{K_2} \right) \tag{2.15}$$

which tends to $-\infty$ as $r \to \infty$. ∎

Note that Assumption 2.A implies the condition of Lemma 6. From this we obtain exponential tightness as follows.

Corollary 7 *If $\lambda_i(\vec{x}) \le K(1+|\vec{x}|)$ for all i then the $\{\vec{z}^n\}$ are exponentially tight. Consequently, it suffices to prove the Large Deviations estimates for bounded sets, uniformly in the initial conditions over compact sets.*

Proof. Given $\alpha > 0$, by Lemma 6 there is a r_α so that

$$\limsup_n \frac{1}{n} \log \mathbb{P}_x \left(\sup_{0\le t\le T} |\vec{z}^n(t)| > r_\alpha \right) \le -\alpha \tag{2.16}$$

for all initial conditions $\vec{x}$ in a bounded set. But then, since the rates are Lipschitz, they are bounded over the region $|\vec{z}^n(t)| \le r_\alpha$. Therefore, by [7, Lemma 5.58] the $\{\vec{z}^n\}$ are exponentially tight. Thus by [4, Lemma 1.2.18(a)] it suffices to prove the upper bound for compact (and in particular bounded) sets. By [7, (5.31–5.32)] it suffices to prove the lower bound on balls: however, all paths in a ball are by definition bounded. ∎

Corollary 8 *Let G be a closed set and $\lambda_i(\vec{x}) \le K(1+|\vec{x}|)$ for all i. Assume that there are $R_k \to \infty$ and open $B_k \subset G^o$ so that the conditions of Corollary 4 hold with B_k replacing B, $G \cap B(0,R_k)$ replacing G and with $\partial G \cap B(0,R_k)$ replacing ∂G. Then Theorems 2, 3 and Corollary 4 hold.*

Note that the constants in Assumptions 1–2 need not be uniform: they may depend on R_k. Since it suffices to derive our results in the case that G is compact, we shall do so since this simplifies the technicalities. Thus for the rest of the paper we restrict the process to a bounded set. We shall therefore assume henceforth that G is bounded and so $\overline{\lambda} \triangleq \max_j \sup_{\vec{x}} \lambda_j(\vec{x}) < \infty$.

Lemma 9 *There exists a continuous monotone "scale function" $s(\delta)$ such that $s(0)=0$, $s(\delta)>0$ for any $\delta>0$, and for any $\vec{x} \in G$ we have*

$$\lambda_i(\vec{x}) \ge s(d(\vec{x}, \partial G)) \text{ for all } i.$$

Proof. Take $s(\delta) \stackrel{\triangle}{=} \inf \{\lambda_i(\vec{x}) \ : \ i; \ \vec{x} : d(\vec{x}, \partial G) \geq \delta\}$ The properties follow by continuity of $\lambda_i(\vec{x})$ and compactness of G. ∎

3 Preliminaries

This section collects some preliminary definitions and estimates. We use an alternative form of ℓ in many of the proofs. Let

$$f(\mu, \lambda) \stackrel{\triangle}{=} \sum_j \lambda_j - \mu_j + \mu_j \log \frac{\mu_j}{\lambda_j} \tag{3.1}$$

$$K_{\vec{y}} \stackrel{\triangle}{=} \left\{ \mu : \mu_j \geq 0, \ \sum_j \mu_j \vec{e}_j = \vec{y} \right\}. \tag{3.2}$$

Note that, by definition, $K_{\vec{y}}$ is non empty if and only if $\vec{y} \in \mathcal{C}_x$.

Lemma 10 *[7, Theorem 5.26] Fix x and y in $\mathbb{R}^d$. Then the two representations are equivalent, that is $\ell(x, y) = \inf \{f(\mu, \lambda(x)) : \mu \in K_{\vec{y}}\}$.*

Note that if $\lambda_j(x) = 0$, then it plays no role in the definition of ℓ. Further, since f is to be minimized, if $\lambda_j(x) = 0$, the corresponding μ_j is necessarily 0, too, so that the j component again plays no role. The lemma also means that either both expressions are finite, or both are infinite.

Lemma 11 *[7, Lemma 5.20] There exists a constant κ such that for all $\vec{x}$ and all $\vec{y} \in \mathcal{C}_{\vec{x}}$ there exists $\vec{a}$ such that*

$$\vec{y} = \sum_{i=1}^{k} a_i \vec{e}_i; \ a_i \geq 0; \ |\vec{a}| \leq \kappa \, |\vec{y}| \, .$$

Lemma 12 *Fix λ. The function f is nonnegative, strictly convex in μ for $\mu_j \geq 0$ with compact level sets. On $\mathcal{C}_x$ the following hold. The function f has a unique minimum over $K_{\vec{y}}$, denoted μ^*. Both the minimum and μ^* are continuous in y. In addition, given B, there are constants C_0, C_1, C_2, D so that for all $|\lambda_i| < B$,*

$$|\mu^*| \leq C_0 \qquad \text{for } |\vec{y}| \leq D \tag{3.3}$$

$$|\mu^*| \leq C_1 |\vec{y}| \quad \text{for } |\vec{y}| > D \tag{3.4}$$

$$|\mu^*| \geq C_2 |\vec{y}|. \tag{3.5}$$

Proof. The function $1 - x + x \log x$ is nonnegative and strictly convex for $x \geq 0$, as seen by differentiation. Level sets are compact since the function grows super-linearly. Existence of a minimum now follows since $K_{\vec{y}}$ is closed. Uniqueness follows from strict convexity since $K_{\vec{y}}$ is a convex set. Continuity follows since f is continuous and superlinear in μ, $K_{\vec{y}}$ is continuous in y, and the *a priori* bound on μ^* Lemma 11. The first bound is a consequence of the same lemma and the compactness of the level sets. The second bound follows from the same argument since f grows faster than linearly in μ. The last bound is immediate from the definition of $K_{\vec{y}}$. ∎

We now establish that $I_{[0,T]}$, as defined in (2.6), is a good rate function.

Theorem 13 *[7, Prop. 5.49 and Cor. 5.50] Assume the $\lambda_i(\vec{x})$ are bounded and continuous. Then, for each $\vec{x}$, $I_{[0,T]}(\cdot)$ is a good rate function under either the sup norm or the Skorohod J_1 metric.*

Proof. Identical to the proof in [7]; the additional assumption (that the $\log \lambda_i$ are bounded) is not used in the proof (the original theorem was stated with unnecessary conditions). ∎

Lemma 14 *[Uniform absolute continuity: [7, Lemma 5.18]] Assume the $\lambda_i(\vec{x})$ are bounded, let $I_{[0,T]}(\vec{r}) \leq K$, and fix some $\epsilon > 0$. Then there is a δ, independent of $\vec{r}$, such that for any collection of nonoverlapping intervals*

$$\{[t_j, s_j],\ j = 1, \ldots, J\} \text{ with } \sum_{j=1}^{J} s_j - t_j = \delta,$$

in $[0, T]$ with total length δ, we have

$$\sum_{j=1}^{J} |\vec{r}(t_j) - \vec{r}(s_j)| < \varepsilon.$$

Lemma 15 *Assume $\lambda_i(\vec{x}) \leq \overline{\lambda}$ for all i and define B_i as in Assumption 1. Then for any $T > 0$, $K > 0$ there is a J such that if $I_{[0,T]}(\vec{r}) \leq K$ then there are $0 = t_0 < t_1 < \ldots < t_J = T$ and j_i so that $\vec{r}(t) \in B_{j_i}$, $t_{i-1} \leq t \leq t_i$.*

Proof. We first claim that there is an $\alpha > 0$ such that for any $\vec{x} \in G$ there is an i such that $B(\vec{x}, \alpha) \subset B_i$. This is easily proved by contradiction; if false, then there is a sequence of points $\vec{x}_j \to \vec{x}$ with diminishing open balls; but the point $\vec{x}$ is contained in the interior of some B_j.

Now for any path $\vec{r}(t)$ with $I_{[0,T]}(\vec{r}) \leq K$, we break up the path according to the following rules. We initially assign the ball B_i to the path at time 0 if $B(\vec{r}(0), \alpha) \subset B_i$ (we may take any B_i that satisfies this constraint). We maintain the choice i until such time that $B(\vec{r}(t), \frac{\alpha}{2}) \not\subset B_i$. At this time, change to any B_j with $B(\vec{r}(0), \alpha) \subset B_j$. By the uniform absolute continuity of $\vec{r}$, Lemma 14, there is a minimum time τ between any change in balls B_i. So $J = \frac{T}{\tau}$ suffices as a bound on the number of pieces. Note that at any switchover time t between balls B_i and B_j we have both $\vec{r}(t) + \frac{\alpha}{4}\vec{v}_i$ and $\vec{r}(t) + \frac{\alpha}{4}\vec{v}_j$ are contained in $B_i \cap B_j$. ∎

4 Upper Bound—Boundary case

We prove the upper bound (2.7) along the same general lines as [7, Theorem 5.54]. There are two difficulties to overcome, both technical. The first is that the previous proof used the fact that shifting a path $\vec{r}(t)$ to $\vec{r}(t) + \delta\vec{v}_i$ resulted in a continuous change in the cost $I_{[0,T]}(\vec{r})$. This is not true in the present case. A shift may result in a path being outside the set G, thus having infinite cost. Another difficulty is that the approximating functional ℓ^δ, used in the proof of the upper bound, is difficult to estimate in the present case. Our solution is to make approximating functionals ℓ^Δ and ℓ^m that are finite for all absolutely continuous paths in G. We then need to calculate how well these functionals approximate ℓ, and to show that the requisite bounds obtain. The main technical estimate we need is that every path $\vec{r}$, having approximating cost $I^\Delta(\vec{r})$, has a path $\vec{r}_1$ at a distance no more than ε, with true cost $I(\vec{r}_1) \approx I^\Delta(\vec{r})$.

Although we shall use Assumptions 1–2, in fact for the upper bound we do not need Lipschitz continuity of the rates. For our proof it suffices that the rates are absolutely continuous (note that Lipschitz functions are absolutely continuous). To emphasize this fact we denote by K_λ the modulus of continuity of the rates. We remind you that $K_\lambda(\delta)$ is the modulus of continuity of all $\lambda_i(\vec{x})$ means $|\lambda_i(\vec{x}) - \lambda_i(\vec{y})| \leq K_\lambda(|\vec{x} - \vec{y}|)$, where $K_\lambda(\delta)$ is continuous and $K_\lambda(0) = 0$.

Lemma 16 *Under assumptions 1 and 2.C there is a constant C such that for every $x \in \partial G$ and $x \in B_i$, $i \leq I_1$, and every η small enough, the cost of the linear path $\vec{r}(t) \stackrel{\triangle}{=} x + t\vec{v}_i$, $0 \leq t \leq \eta$ is less than $C\eta$. In other words, with $T = \eta$,*

$$I_{[0,T]}(\vec{r}) \leq C\eta. \tag{4.1}$$

Proof. By Assumption 2.C, we can find a set of $\vec{e}_i$ and a_i with $\sum_i a_i \vec{e}_i = \vec{v}$ and with $\lambda_i(\vec{r}(t)) > \gamma/2 > 0$ for all $t \leq \eta$. The a_i are bounded since $\vec{v}$ is, by Lemma 3. So, by the representation of ℓ in Lemma 10, taking the $\mu_i = a_i$ and λ_i as given yields a bound on ℓ. ■

The proof of the upper bound of [7] uses an "approximate" rate function $\overset{\delta}{I}$: here we require several different approximations. The functional I^Δ is defined via the function g^Δ as follows:

$$g^\Delta(\vec{x}, \vec{\theta}) \triangleq \sum_{i=1}^{k} \sup_{\vec{z}_i \in B_\delta(\vec{x})} \left[\lambda_i(\vec{z}_i) \left(e^{\langle \vec{\theta}, \vec{e}_i \rangle} - 1 \right) \right] \tag{4.2}$$

$$\ell^\Delta(\vec{x}, \vec{y}) \triangleq \sup_{\vec{\theta} \in \mathbb{R}^d} \left(\langle \vec{\theta}, \vec{y} \rangle - g^\Delta(\vec{x}, \vec{\theta}) \right) \tag{4.3}$$

and $I^\Delta_{[0,T]}(\vec{r})$ is now defined like $I_{[0,T]}(\vec{r})$, with ℓ^Δ replacing ℓ. The functional $\overset{\delta}{I}$ was defined in a similar manner [7, Defs. 5.36–38], but starting with

$$g^\delta(\vec{x}, \vec{\theta}) \triangleq \sup_{\vec{z}_i \in B_\delta(\vec{x})} \sum_{i=1}^{k} \lambda_i(\vec{z}_i) \left(e^{\langle \vec{\theta}, \vec{e}_i \rangle} - 1 \right). \tag{4.4}$$

The difference between g^δ and g^Δ is that the supremum and sum are interchanged. In g^Δ there are many different $\vec{z}_i$ where each maximum is attained; in g^δ there is only one $\vec{z}$. We are abusing notation a bit here; the terms labeled with a $^\Delta$ are defined in terms of a parameter δ. Of course, we are supposing that the value of Δ is δ, but the difference in notation should make clear what we mean. Clearly we have $g^\Delta(\vec{x}, \vec{\theta}) \geq g^\delta(\vec{x}, \vec{\theta})$ for every δ, $\vec{x}$, and $\vec{\theta}$, so $\ell^\Delta(\vec{x}, \vec{y}) \leq \ell^\delta(\vec{x}, \vec{y})$ and $I^\Delta_{[0,T]}(\vec{r}) \leq \overset{\delta}{I}_{[0,T]}(\vec{r})$. We could have avoided the use of $\overset{\delta}{I}$ entirely in this paper, except that there are many bounds available for it already, so instead of having to prove all these bounds for I^Δ separately, we will use them in conjunction with the simple relationship between $\overset{\delta}{I}$ and I^Δ.

The advantage of using I^Δ instead of $\overset{\delta}{I}$ is that I^Δ has a representation as a change of measure as follows. Note that since λ_i is continuous and since the supremum is taken for each i separately, $g^\Delta(\vec{x}, \vec{\theta})$ is the supremum over a convex set of $\{\lambda_i\}$ of a function which is linear in the λ_i. Therefore

$$\ell^\Delta(\vec{x},\vec{y}) \stackrel{\triangle}{=} \sup_{\vec{\theta}\in\mathbb{R}^d} \inf_{\vec{z}_i\in B_\delta(\vec{x})} \left(\langle\vec{\theta},\vec{y}\rangle - \sum_{i=1}^{k} \lambda_i(\vec{z}_i) \left(e^{\langle\vec{\theta},\vec{e}_i\rangle} - 1\right)\right) \tag{4.5}$$

$$= \inf_{\vec{z}_i\in B_\delta(\vec{x})} \sup_{\vec{\theta}\in\mathbb{R}^d} \left(\langle\vec{\theta},\vec{y}\rangle - \sum_{i=1}^{k} \lambda_i(\vec{z}_i) \left(e^{\langle\vec{\theta},\vec{e}_i\rangle} - 1\right)\right) \tag{4.6}$$

$$= \sup_{\vec{\theta}\in\mathbb{R}^d} \left(\langle\vec{\theta},\vec{y}\rangle - \sum_{i=1}^{k} \lambda_i(\vec{z}_i^*) \left(e^{\langle\vec{\theta},\vec{e}_i\rangle} - 1\right)\right) \tag{4.7}$$

$$= \sum_{i=1}^{k} \left(\lambda_i(\vec{z}_i^*) - \mu_i + \mu_i \log\frac{\mu_i}{\lambda_i(\vec{z}_i^*)} \right), \tag{4.8}$$

where the $\vec{z}_i^* \in B_\delta(\vec{x})$. Equation (4.6) follows from the min-max saddle point theorem, since the function in the definition of $\ell^\Delta(\vec{x},\vec{y})$ is concave in $\vec{\theta}$ and linear, hence convex in the λ_i; furthermore, the infimum is taken over a bounded convex set of λ_i. Equation (4.7) follows by taking a minimizing $\vec{z}_i$, and equation (4.8) follows from the representation theorem for $\ell(\vec{x},\vec{y})$.

For technical complexity, but to make easy proofs, we also define a function ℓ^m as follows. Define $\vec{\lambda}^m$ by

$$\lambda_j^m(\vec{x}) \stackrel{\triangle}{=} \max\{\lambda_j(\vec{x}), 1/m\}. \tag{4.9}$$

Define ℓ^m and I^m through (2.5)–(2.6) but using he modified rates $\lambda_j^m(\vec{x})$. To complete this sequence of definitions, we define the set of cheap paths in any of the metrics:

$$\Phi_{\vec{x}}(K) \stackrel{\triangle}{=} \{\vec{r}(t) : \vec{r}(0) = \vec{x}, I_{[0,T]}(\vec{r}) \le K\} \tag{4.10}$$

$$\Phi_{\vec{x}}^\Delta(K) \stackrel{\triangle}{=} \{\vec{r}(t) : \vec{r}(0) = \vec{x}, I_{[0,T]}^\Delta(\vec{r}) \le K\} \tag{4.11}$$

$$\Phi_{\vec{x}}^m(K) \stackrel{\triangle}{=} \{\vec{r}(t) : \vec{r}(0) = \vec{x}, I_{[0,T]}^m(\vec{r}) \le K\}. \tag{4.12}$$

Lemma 17 *Assume 1–2. Fix i. For each $T > 0$, $K > 0$ and $0 < \varepsilon < K$ there is an $0 < \eta_0 < \varepsilon$ such that for any $0 < \eta < \eta_0$ there is a m_0 with the following property. If $m > m_0$ and if the path $\vec{r}(t)$ takes values in B_i with $I_{[0,T]}^m(\vec{r}) < K - \varepsilon$, then the path $\vec{r}_2(t) \stackrel{\triangle}{=} \vec{r}(t) + \eta\vec{v}_i$ satisfies $I_{[0,T]}(\vec{r}_2) < K$.*

Proof. Let $\vec{v} = \vec{v}_i$,

$$\ell_1(t) = \ell^m(\vec{r}(t), \vec{r}\,'(t)) \tag{4.13}$$

$$\ell_2(t) = \ell(\vec{r}(t) + \eta\vec{v}, \vec{r}\,'(t)) \tag{4.14}$$

$$\lambda_i^1(t) = \lambda_i^m(\vec{r}(t)) \tag{4.15}$$

$$\lambda_i^2(t) = \lambda_i(\vec{r}(t) + \eta\vec{v}) \tag{4.16}$$

$$\mu_i^*(t) = \mu_i^{m,*}(\vec{r}(t), \vec{r}\,'(t)) \tag{4.17}$$

where $\mu^{m,*}$ is optimal (see Lemma 12) for jump rates λ_i^m. We denote by $\lambda^1(t)$ the vector with coordinates $\lambda_i^1(t)$ and similarly for λ^2 and μ^*. By Lemma 10,

$$\ell_2(t) \leq f\left(\mu^*(t), \lambda_i^2(t)\right).$$

Choose $m_0 > 1/\eta$. Then by continuity of λ_i

$$|\lambda_i^2(t) - \lambda_i^1(t)| \leq \frac{1}{m} + |\lambda_i^2(t) - \lambda_i(\vec{r}(t))| \leq K_\lambda(\eta) + \eta \triangleq K'_\lambda(\eta).$$

Therefore

$$\ell_2(t) - \ell_1(t) \leq \sum_{i=1}^k \lambda_i^2(t) - \lambda_i^1(t) + \mu_i^*(t)\log\frac{\lambda_i^1(t)}{\lambda_i^2(t)} \tag{4.18}$$

$$\leq k \cdot K'_\lambda(\eta) + \sum_{i=1}^k \mu_i^*(t)\log\frac{\lambda_i^1(t)}{\lambda_i^2(t)}. \tag{4.19}$$

By Assumption 2.C there exists a $\gamma > 0$ such that if $\lambda_i(\vec{x}) < \gamma$ then $\lambda_i(\vec{x} + \eta\vec{v}) \geq \lambda_i(\vec{x})$. Increase m_0 so that (recall $s(\eta)$ was defined in Lemma 9)

$$m_0 > \max\{1/\gamma, 1/s(\eta)\}.$$

For the rest of the proof let $m > m_0$. Then if $\lambda_i(\vec{r}(t)) < \gamma$ we have $\lambda_i^1(t) < \gamma$ and $\lambda_i^2(t) \geq \lambda_i^1(t)$. Therefore if $\lambda_i(\vec{r}(t)) < \gamma$ then $\log \lambda_i^1(t)/\lambda_i^2(t) < 0$. If however $\lambda_i(\vec{r}(t)) > \gamma$ then $\lambda_i(\vec{r}(t)) = \lambda_i^1(t)$, and so, with $\vec{x} = \vec{r}(t)$,

$$\log\frac{\lambda_i^1(t)}{\lambda_i^2(t)} = \log\frac{\lambda_i(\vec{x})}{\lambda_i(\vec{x} + \eta\vec{v})} \tag{4.20}$$

$$\leq \log\frac{\lambda_i(\vec{x})}{\lambda_i(\vec{x}) - K_\lambda(\eta)} \tag{4.21}$$

$$\leq \log\frac{\gamma}{\gamma - K_\lambda(\eta)} \tag{4.22}$$

$$= \log\frac{1}{1 - \frac{K_\lambda(\eta)}{\gamma}} \tag{4.23}$$

$$\leq \frac{2K_\lambda(\eta)}{\gamma} \tag{4.24}$$

for $K_\lambda(\eta) < \frac{\gamma}{2}$, since $\log\frac{1}{1-x} < 2x$ for $0 < x \le 1/2$. By [7, Lemma 5.17] and Lemma 12 there exists constants C_1, C_3, B_1 such that if $|\vec{y}| > B_1$ then $\ell(\vec{x},\vec{y}) \ge C_3|\vec{y}|\log|\vec{y}|$ and and $|\mu^*| \le C_1|\vec{y}|$. So if $|\vec{r}\,'(t)| \ge B_1$ then using (4.18)–(4.24)

$$\ell_2(t) - \ell_1(t) \le kK'_\lambda(\eta) + C_1|\vec{r}\,'(t)|\frac{2K_\lambda(\eta)}{\gamma}.$$

But by the estimate above, if $|\vec{r}\,'(t)| > B_1$ then $|\vec{r}\,'(t)| < \frac{\ell_1(t)}{C_3\log|\vec{r}\,'(t)|}$. Thus if $|\vec{r}\,'(t)| > B_1$ we get

$$\ell_2(t) - \ell_1(t) \le kK'_\lambda(\eta) + C\frac{\ell_1(t)}{C_3\log|\vec{r}\,'(t)|}\frac{2K_\lambda(\eta)}{\gamma} \tag{4.25}$$

$$\le kK'_\lambda(\eta) + C\frac{\ell_1(t)}{C_3\log B_1}\frac{2K_\lambda(\eta)}{\gamma}. \tag{4.26}$$

Now consider the case $|\vec{r}\,'(t)| \le B_1$. By Lemma 12, $|\mu^*(t)| \le K_\mu B_1$ for some K_μ. Therefore

$$\ell_2(t) - \ell_1(t) \le kK'_\lambda(\eta) + K_\mu B_1\frac{2K_\lambda(\eta)}{\gamma}.$$

Putting together the estimates for large and small values of $\vec{r}\,'(t)$ we obtain

$$\ell_2(t) - \ell_1(t) \le \left(kK'_\lambda(\eta) + K_\mu B_1\frac{2K_\lambda(\eta)}{\gamma} + C\frac{\ell_1(t)}{C_1\log B_1}\frac{2K_\lambda(\eta)}{\gamma}\right). \tag{4.27}$$

Therefore if $I^m_{[0,T]}(\vec{r}\,) \le K - \varepsilon$ then

$$I_{[0,T]}(\vec{r} + \eta\vec{v}) \le K - \varepsilon + T\left(kK'_\lambda(\eta) + K_\mu B_1\frac{2K_\lambda(\eta)}{\gamma} + \frac{C(K-\varepsilon)2K_\lambda(\eta)}{TC_1\log B_1\gamma}\right). \tag{4.28}$$

Thus for small enough η, $I_{[0,T]}(\vec{r} + \eta\vec{v}) \le K$. ∎

Lemma 18 *Assume 1–2. Let A be a compact set in G. For each $T > 0$, $K > 0$ and $0 < \varepsilon < K$ there is an $0 < \eta_0 < \varepsilon$ such that for any $0 < \eta < \eta_0$ there is a m_0 with the following property. For any $x \in A$, $m > m_0$, and any path $\vec{r}(t)$ with $\vec{r}(0) = \vec{x}$ and $I^m_{[0,T]}(\vec{r}\,) < K - \varepsilon$, there is a path $\vec{r}_2(t)$ with $\vec{r}(0) = \vec{x}$ with $I_{[0,T]}(\vec{r}_2) < K$ and $d(\vec{r}, \vec{r}_2) \le \varepsilon$.*

This lemma corresponds to [7, Lemma 5.48]. It is based on a direct construction, having the path $\vec{r}_2$ composed of a number of segments. The main segments are made of shifted pieces of $\vec{r}$, using Lemma 17 to estimate the costs of the segments, with the initial part of the path estimated by Lemma 16, and the segments stitched together with the aid of Lemma 16 as well.

Proof. Given $\vec{r}(t)$ with $\vec{r}(0) = \vec{x}$ and $I^m_{[0,T]}(\vec{r}) < K - \varepsilon$, take J as defined in Lemma 15. Recall from Assumption 1.(iii), there is a number $\beta > 0$ such that for any $\vec{x} \in B_i$ where B_i is one of the boundary neighborhoods, then for small ε we have $d(\vec{x} + \varepsilon\vec{v}_i, \partial G) \geq \beta\varepsilon$. Clearly $\beta \leq 1$.

We now construct the path $\vec{r}_2$ from $\vec{r}$ using a parameter η that will be chosen later. Let $0 = t_0, t_1, \ldots, t_J$ represent the switchover times of $\vec{r}(t)$, as in Lemma 15. Recall the definition of α from the proof of Lemma 15. If $B(\vec{r}(0), \alpha) \subset B_i$ where B_i is one of the boundary neighborhoods ($i \leq I_1$, cf. Assumption 1.(i)), then take $\vec{r}_2(t) = \vec{r}(0) + \vec{v}_i t$ for $0 \leq t \leq \eta$. For $\eta \leq t \leq t_1 + \eta$ let $\vec{r}_2(t) = \vec{r}(t - \eta) + \eta\vec{v}_i$. Then for $t_1 + \eta \leq t \leq t_1 + \eta + 3\eta/\beta$ let

$$\vec{r}_2(t) = \vec{r}(t_1) + \eta\vec{v}_i + (t - t_1 - \eta)\vec{v}_j; \tag{4.29}$$

here j is the index of the neighborhood B_j that $\vec{r}(t)$ switches to at time t_1, as defined in Lemma 15. Then from time $t_1 + \eta + 3\eta/\beta$ until time $t_2 + \eta + 3\eta/\beta$ we let $\vec{r}_2(t) = \vec{r}(t - \eta - 3\eta\beta) + \eta\vec{v}_i + (3\eta/\beta)\vec{v}_j$. We continue in this fashion, with switchover k having $\vec{r}_2(t)$ following a linear path of length $\eta(3/\beta)^{k-1}$ in direction $\vec{v}_k$, followed by a segment parallel to $\vec{r}(t)$. Note that we need not include the paths of length $\eta(3/\beta)^{k-1}$ in direction $\vec{v}_k$ if B_k is not a boundary neighborhood.

There are three things to check about this path. First, if η is small enough, does $\vec{r}_2(t)$ stay within ε of $\vec{r}(t)$? Second, does $\vec{r}_2$ remain in G? Third, does it satisfy $I_{[0,T]}(\vec{r}_2) \leq K$? If all three items hold, then the lemma will be proved. These estimates are hardest if all the paths of type $\eta(3/\beta)^{k-1}$ in direction $\vec{v}_k$ are included, so we assume without loss of generality that they are.

The first thing is easy to verify. Since $I^m_{[0,T]}(\vec{r}) < K - \varepsilon$ by assumption, $\vec{r}(t)$ is uniformly absolutely continuous, so by choosing η small enough, we can ensure that the time shifts introduced in the definition of $\vec{r}_2$ don't cause the paths to differ by more than $\varepsilon/2$. Furthermore, the number of segments J is bounded, so the difference introduced in the segments $\eta(3/\beta)^k$ adds up to less than $\varepsilon/2$ if η is chosen small enough.

For the second point, this is the reason we chose $3/\beta$ as a multiplier. By definition of β, the point $\vec{r}_2 + \eta(3/\beta)^k\vec{v}_k$ is at least $\beta\eta(3/\beta)^k$ from ∂G. Since $3/\beta > 3$, the sum of all the previous shifts has total length less than $(\eta\beta/2)(3/\beta)^k$. Therefore the point is at least $(\eta\beta/2)(3/\beta)^k$ from ∂G when the kth shift is finished. The beginning of the shift also occurs in the interior of G by construction, whenever the total shifts are of length less than $\alpha/4$. This demonstrates that $\vec{r}_2 \in G$ when η is small enough.

For the third point, we assume that η has been chosen small enough to satisfy the first two points, and note that, after the initial time η, the path

$\vec{r}_2(t)$ remains at least $\beta\eta/2$ away from ∂G. By Lemma 16, there is a uniform constant C' such that for $0 \le t \le \eta$

$$I_{[0,t]}(\vec{r}_2) \le C't. \tag{4.30}$$

We use this estimate on every $\eta(3/\beta)^k$ path; the total cost is linear in η, so can be made less than $\varepsilon/2$. Lemma 17 enables us to bound the cost of each other segment of $\vec{r}_2$ uniformly as no more than $\varepsilon_1(\eta)$ plus the I^m-cost of the corresponding segment $\vec{r}$, where $\varepsilon_1(\eta)$ goes to 0 with η. Choose η small enough so that $\varepsilon_1(\eta) \le \varepsilon/2J$. Then choose m large enough that Lemma 17 applies. Then the total additional cost is bounded by $\varepsilon/2$ for the segments of $\vec{r}_2$. This concludes the estimate, and hence the proof. ∎

Lemma 19 *Under the assumption of bounded continuous λ_i, given $\varepsilon > 0$ there exists an $m_0 > 0$ such that for all positive $m > m_0$ there exists a $\delta_0 > 0$ such that for all $\delta < \delta_0$,*

$$\ell^m(\vec{x},\vec{y}) \le \varepsilon + (1+\varepsilon)\ell^\Delta(\vec{x},\vec{y}). \tag{4.31}$$

Proof. We consider separately the cases $|\vec{y}| > B$ and $|\vec{y}| \le B$. Here we take B so that $\ell(\vec{x},\vec{y}) > CB\log B$ for all $\vec{x}$; this is possible by [7, Lemma 5.17].

If $|\vec{y}| > B$ then, using (4.8) and the reasoning leading to Equation (4.24),

$$\ell^m(\vec{x},\vec{y}) - \ell^\Delta(\vec{x},\vec{y}) \le K_\lambda(\delta)\left(k + \frac{2C\ell^\Delta(\vec{x},\vec{y})}{\gamma C_1 \log B}\right) \tag{4.32}$$

Therefore, by choosing δ small enough, we can make the right-hand side of (4.32) smaller than $\varepsilon/3$.

Now if $|\vec{y}| \le B$, take μ as the optimal jump rate for $\lambda^\Delta(\vec{x})$. We have $\mu_i \le CB$ for all i. Then

$$\ell^m(\vec{x},\vec{y}) - \ell^\Delta(\vec{x},\vec{y}) \le \sum_i \lambda_i^m(\vec{x}) - \lambda_i^\Delta(\vec{x}) + \mu_i \log\frac{\lambda_i^\Delta(\vec{x})}{\lambda_i^m(\vec{x})} \tag{4.33}$$

$$\le k(K_\lambda(\delta) + 1/m) + \sum_i \mu_i \log\frac{\lambda_i^\Delta(\vec{x})}{\lambda_i^m(\vec{x})}. \tag{4.34}$$

If $\lambda_i(\vec{x}) \le 1/m - K_\lambda(\delta)$ then $\lambda_i^\Delta(\vec{x}) \le 1/m$, so, since $\lambda_i^m(\vec{x}) \ge 1/m$ for each $\vec{x}$, $\log\frac{\lambda_i^\Delta(\vec{x})}{\lambda_i^m(\vec{x})} \le 0$. Furthermore, if $\lambda_i(\vec{x}) > 1/m - K_\lambda(\delta)$, then

$$\log\frac{\lambda_i^\Delta(\vec{x})}{\lambda_i^m(\vec{x})} \le \log\frac{1/m + K_\lambda(\delta)}{1/m} \le mK_\lambda(\delta). \tag{4.35}$$

Hence, by choosing δ small enough that $kmK_\lambda(\delta) < \varepsilon/3$, we obtain, for $|\vec{y}| \le B$, $\ell^m(\vec{x},\vec{y}) - \ell^\Delta(\vec{x},\vec{y}) < \varepsilon/3$. Combined with the result of the previous paragraph, this finishes the proof. ∎

Corollary 20 *Assume the λ_i are bounded and continuous. Given ε, K, and $T > 0$ there exists an $m_0 > 0$ such that for all positive $m > m_0$ there exists a $\delta_0 > 0$ so that $\delta < \delta_0$ and $I^\Delta_{[0,T]}(\vec{r}) \le K - \varepsilon$ imply $I^m_{[0,T]}(\vec{r}) \le I^\Delta_{[0,T]}(r) + \varepsilon$.*

Corollary 20 and Lemma 18 combine to give the following:

Corollary 21 *Assume 1–2. Given $K > 0$ and $\varepsilon > 0$, there exists a $\delta > 0$ such that*

$$\Phi^\Delta_{\vec{x}}(K-\varepsilon) \subset \{\vec{r} : d(\vec{r}, \Phi_{\vec{x}}(K)) \le \varepsilon\}. \tag{4.36}$$

The proof of this corollary is immediate, by using $\varepsilon/2$ to replace ε, and choosing a large enough m. (Specifically, if $\vec{r} \in \Phi^\Delta_{\vec{x}}(K-\varepsilon)$ then, by Corollary 20, $\vec{r} \in \Phi^m_{\vec{x}}(K-\varepsilon/2)$. By Lemma 18, there is an $\vec{r}_2$ within $\varepsilon/2$ of $\vec{r}$ satisfying $I_{[0,T]}(\vec{r}_2) \le K$.) ∎

Corollary 21 shows that, by choosing δ and η small enough,

$$d(\vec{r}, \Phi_{\vec{y}}(K-4\varepsilon)) > \eta/2 \text{ implies } d(\vec{r}, \Phi^\Delta_{\vec{y}}(K-4\varepsilon-\eta/4)) > \eta/4. \tag{4.37}$$

We now state two technical lemmas in measure theory that are used in the proof of Lemma 24. Lemma 22 is used only for the proof of Lemma 23. We let $\mathrm{Leb}(A)$ denote the Lebesgue measure of a set A.

Lemma 22 *Let $u(t)$ be a nonnegative, absolutely continuous function on $[0,T]$. Then given $\delta > 0$ there exists an $\eta > 0$, a set $A \subset [0,T]$, and a finite collection $\{C_i\}$ of intervals so that $\mathrm{Leb}(A) < \delta$ and, for each i, either $\inf\{u(t) : t \in C_i > \eta\}$ or $u(t) = 0$ for all $t \in C_i \setminus A$.*

Proof. If $u(0) > 0$ set $t_1 \stackrel{\triangle}{=} \inf\{t > 0 : u(t) = 0\}$. Then by continuity $\inf\{u(t) : 0 \le t \le t_1 - \delta/2\} > 0$, and so it suffices to establish the result when $u(0) = 0$. By a similar argument we may assume $u(T) = 0$.
Given t, if $u(t) > 0$ then by continuity there exists an open interval O_t containing t so that $u(s) \in O_t$ for all $s \in O_t$ and $u(s) \to 0$ as $s \to \partial O_t$. Let $m_t \stackrel{\triangle}{=} \sup\{u(s) : s \in O_t\}$. Since u is absolutely continuous, there is a finite number of disjoint open intervals with $m_t > 1/m$. Therefore there exists a

countable collection $\{O_i\}$ of disjoint open intervals so that $u(t) > 0$ if and only if $t \in O_i$ for some i. Fix N large so that

$$\mathrm{Leb}\left\{\cup_{N+1}^{\infty} O_i\right\} < \frac{\delta}{2}\ . \tag{4.38}$$

For $i \leq N$ let $\tilde{C}_i \subset O_i$ be a closed interval such that

$$\mathrm{Leb}\{O_i \setminus \tilde{C}_i\} \leq \frac{\delta}{2N+1}.$$

Let C_i be the finite collection of closed intervals that cover $[0,T] \setminus \cup_i \tilde{C}_i$ and let

$$A \triangleq \left\{\cup_{N+1}^{\infty} O_i\right\} \cup \left\{\cup_{i=1}^{N} O_i \setminus \tilde{C}_i\right\}\ .$$

Then by construction, $\mathrm{Leb}(A) < \delta$ and $u(t) = 0$ on $t \in C_i \setminus A$. Since there are only a finite number of $\tilde{C}_i$, we obtain by continuity $\inf\{u(t) : t \in \tilde{C}_i$ for some $i\} > \eta > 0$. ∎

For vectors $\vec{\theta}, \vec{\lambda}$ and $\vec{y}$ in $\mathbb{R}^d$ define

$$\ell(\vec{\theta}, \vec{\lambda}, \vec{y}) \triangleq \left(\langle\vec{\theta}, \vec{y}\rangle - \sum_j \lambda_j \left(e^{\langle\vec{\theta}, \vec{e}_j\rangle} - 1\right)\right). \tag{4.39}$$

Our next result extends [7, Lemma 5.43] to the case of rates which are not bounded below, but under the assumption that they are absolutely continuous.

Lemma 23 *Assume that the $\lambda_j(\vec{x})$ are bounded and absolutely continuous. Then for any $\vec{r}$ with $I_{[0,T]}(\vec{r}) < \infty$ and any $\varepsilon > 0$ there exists a step function $\vec{\theta}$ so that*

$$\int_0^T \ell(\vec{\theta}(t), \vec{\lambda}(\vec{r}(t)), \vec{r}\,'(t))\, dt \geq I_{[0,T]}(\vec{r}) - \varepsilon.$$

Proof. Since by definition

$$\ell(\vec{\theta}(t), \vec{\lambda}(\vec{r}(t)), \vec{r}\,'(t)) \leq \ell(\vec{r}(t), \vec{r}\,'(t)),$$

it suffices to establish the result outside a set of small measure: this approximation and the extension of the step function over this set are derived in the proof of [7, Lemma 5.43].

We claim that given δ, there is a partition of $[0, T]$ into a finite collection of intervals C_i and a set A with $\mathrm{Leb}(A) < \delta$ so that the lemma holds on each $C_i \setminus A$. Then the result follows by patching together the step function and dealing with A as above. In the rest of the proof we establish this claim.

Fix j and apply Lemma 22 to the absolutely continuous function $\lambda_j(\vec{r}(t))$ with δ/k. We then obtain a collection $\{C_i^j,\ A^j,\ 1 \le j \le k,\ 1 \le i \le N\}$ so that either $\lambda_j(\vec{r}(t)) = 0$ for all $t \in C_i^j \setminus A^j$ or $\lambda_j(\vec{r}(t)) > \eta_j > 0$ for all $t \in C_i^j$. Set $A \stackrel{\triangle}{=} \cup_j A^j$ and $\eta = \min_j \eta_j$. Then $\mathrm{Leb}(A) < \delta$ and $\eta > 0$. Fix a subset $\alpha \subset \{1, \ldots, k\}$. Using intersections of the sets C_i^j we obtain a finite collection of intervals C_i^α with the following properties:

$$\lambda_j(\vec{r}(t)) = 0, \quad t \in C_i^\alpha \setminus A \text{ for all } j \in \alpha \tag{4.40}$$

$$\lambda_j(\vec{r}(t)) > \eta, \quad t \in C_i^\alpha \text{ for all } j \notin \alpha. \tag{4.41}$$

In particular, $\lambda_j(\vec{r}(t)) > \eta$ for $t \in C_i^\emptyset$. Now fix i and α and consider the process on C_i^α with rates and jump directions $\{\lambda_j(\vec{x}), \vec{e}_j,\ j \notin \alpha\}$. For this process, the assumptions of [7, Lemma 5.43] hold. Moreover, if we denote the local rate function for this process by ℓ^α then, by definition, $\ell^\alpha(\vec{r}(t), \vec{r}\,'(t)) = \ell(\vec{r}(t), \vec{r}\,'(t))$ for $t \in C_i^\alpha \setminus A$. Thus our claim is established and the proof is concluded. ∎

Now we state and prove that the analogue of [7, Proposition 5.62] holds for I^Δ. By [7, Lemma 5.61], under our continuity assumptions, if C is a compact set in $\mathbb{R}^d$, and $\vec{\theta}(t)$ is a fixed step function then for any $\delta > 0$ and compact set $\mathcal{K} \subset \mathcal{K}(M)$ of functions $\vec{r}(t)$ with $\vec{r}(0) \in C$,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_{\vec{x}}(\vec{y}_n \in \mathcal{K}_{\vec{x}}) \le -\inf_{\vec{r} \in \mathcal{K}_{\vec{x}}} \overset{\delta}{I}_{[0,T]}(\vec{r}, \vec{\theta})$$

where $\mathcal{K}_{\vec{x}} \stackrel{\triangle}{=} \{\vec{r} \in \mathcal{K} : \vec{r}(0) = \vec{x}\}$. Note that $I^\Delta > \overset{\delta}{I}$; therefore, we have, under the same assumptions, that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_{\vec{x}}(\vec{y}_n \in \mathcal{K}_{\vec{x}}) \le -\inf_{\vec{r} \in \mathcal{K}_{\vec{x}}} I^\Delta_{[0,T]}(\vec{r}, \vec{\theta}). \tag{4.42}$$

Lemma 24 *Assume the $\lambda_i(\vec{x})$ are bounded and absolutely continuous, and C is a compact set in $\mathbb{R}^d$. Then for each $K > 0$, $\delta > 0$, and $\varepsilon > 0$,*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_{\vec{x}} \left(d\left(\vec{y}_n, \Phi^\Delta_{\vec{x}}(K)\right) > \varepsilon \right) \le -(K - \varepsilon)$$

uniformly in $\vec{x} \in C$.

The proof of this lemma follows exactly the proof of Lemma 5.62 as presented in the book, with the following two exceptions. First, the step function $\vec{\theta}(t)$ is the one defined in Lemma 23. Second, every superscript δ is replaced by the corresponding Δ. This is obvious throughout the proof, with the help of (4.42). ∎

We now state and prove the large deviations upper bound for our process. Recall that for a set F of paths, $I_{\vec{x}}(F) = \inf(I_{[0,T]}(\vec{r}) : \vec{r}(0) = \vec{x}, \vec{r} \in F)$.

Theorem 25 *Assume 1–2. Let F be a closed set in $(D^d[0,T], d_d)$, and let $\vec{x}$ be a point in G. Then*

$$\lim_{\vec{y}\to\vec{x}} \sup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_{\vec{y}}(\vec{z}_n \in F) \leq -I_{\vec{x}}(F) \tag{4.43}$$

where the points $\vec{y}$ are in $F \cap G$.

Proof. Fix a closed set F, and let $K = I_{\vec{x}}(F)$. Suppose for now that $K < \infty$; the same proof will work for $K = \infty$, but some arguments need minor modifications. By [7, Lemma 5.63], for given ε there is a δ so that if $|x - y| < \delta$, $I_{\vec{y}}(F) \geq K - \varepsilon$.

Let

$$C = \cup_{y \in B_\delta(\vec{x}) \cap G} \{\Phi_{\vec{y}}(K - 4\varepsilon)\} \tag{4.44}$$

$$F_\delta = \{\vec{r} \in F : |\vec{r}(0) - \vec{x}| \leq \delta\}. \tag{4.45}$$

Then C is a compact set. Note that $C \cap F_\delta = \emptyset$. By [7, Theorem A.19] there is a number $\eta > 0$ such that $d(C, F_\delta) = \eta$.

Define the random path $\vec{y}_n(t)$ as the linear interpolation of $\vec{z}_n(t)$ with time spacing T/n; that is, at times jT/n for $j = 0, \ldots, n$, $\vec{y}_n(t) = \vec{z}_n(t)$, and $\vec{y}_n(t)$ is linear in between these points. [7, Lemma 5.57] states that, for each $\varepsilon > 0$, uniformly in $\vec{x}$ in bounded sets,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_{\vec{x}}(d(\vec{z}_n, \vec{y}_n) > \varepsilon) = -\infty. \tag{4.46}$$

Now

$$\mathbb{P}_{\vec{y}}(\vec{z}_n \in F) = \mathbb{P}_{\vec{y}}(\vec{z}_n \in F_\delta) \tag{4.47}$$

$$\leq \mathbb{P}_{\vec{y}}(d(\vec{y}_n, F_\delta) < \eta/2) + \mathbb{P}_{\vec{y}}(d(\vec{y}_n, \vec{z}_n) \geq \eta/2). \tag{4.48}$$

Equation (4.46) shows that the second term on the right-hand side of this inequality is negligible.

If $\vec{r}(0) = \vec{y}$, then by definition of η,

$$d(\vec{r}, F_\delta) < \eta/2 \text{ implies } d(\vec{r}, \Phi_{\vec{y}}(K - 4\varepsilon)) > \eta/2 \tag{4.49}$$

Therefore, by equation (4.37) and Lemma 24,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_{\vec{y}}(d(\vec{y}_n, F_\delta) < \eta/2) \tag{4.50}$$

$$\leq \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_{\vec{y}}\left(d\left(\vec{y}_n, \Phi^{\Delta}(K - 4\varepsilon - \eta/4)\right) \geq \eta/4\right) \tag{4.51}$$

$$\leq -(K - 4\varepsilon - \eta/4) \tag{4.52}$$

uniformly in $\vec{y}$ in a compact set. Therefore

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_{\vec{y}}(\vec{z}_n \in F) \leq -(K - 4\varepsilon - \eta/4)$$

whenever $|\vec{y} - \vec{x}| \leq \delta$. Since ε and η can be made arbitrarily small, the theorem is proved. ∎

5 Lower Bound—Boundary Case

Our approach to the proof of the lower bound, Equation (2.8), is mainly standard. Every path $\vec{r}$ in an open set O can be surrounded by a "sausage," a neighborhood of $\vec{r}$, that is entirely contained in O. If we can show that the probability that $\vec{z}_n$ lies in this sausage is about $\exp(-nI_{[0,T]}(\vec{r}))$, then the lower bound will be proved, for if we take a sequence of $\vec{r}$ whose I functions are approximately minimal in O, then we find that the probability that $\vec{z}_n \in O$ is at least the probability that $\vec{z}_n \in$ the sausage around $\vec{r}$, which is about $\exp(-nI_{[0,T]}(\vec{r}))$.

The novelty in the proof is a twofold estimate. The first step is to show that, for any path $\vec{r}$ that lies entirely in a single neighborhood B_i (see Assumption 1.(i)), the probability that $\vec{z}_n$ is near $\vec{r}$ is approximately $\exp(-nI_{[0,T]}(\vec{r}))$. This is done by showing that the path $\vec{r} + \delta\vec{v}_i$, which lies strictly away from the boundary, has rate function $I(\vec{r} + \delta\vec{v}_i) \approx I(\vec{r})$. Then, since the $\lambda_j(\vec{x})$ are bounded away from zero on this path, existing lower bound theory shows that the probability of $\vec{z}_n$ being near this new path is at least $\exp(-nI(\vec{r} + \delta\vec{v}_i))$. Using $I(\vec{r} + \delta\vec{v}_i) \approx I(\vec{r})$ proves the result for such paths $\vec{r}$. The second step is to show that every path $\vec{r}$ with finite cost $I_{[0,T]}(\vec{r})$ can be decomposed into a finite number of pieces $\vec{r}_j$, each of which lies entirely within a ball B_j, and that the endpoints of the shifted pieces $\vec{r}_j + \delta\vec{v}_j$ can be connected with asymptotically negligible cost. This is mostly the same as Lemma 18.

Lemma 26 *Let Assumptions 1–2 hold. Suppose $\vec{r}(t)$ is a path contained in a single B_i with $I_{[0,T]}(\vec{r}) < \infty$. Let $\vec{r}_\delta(t) = \vec{r}(t) + \delta\vec{v}_i$, where $\vec{v}_i$ is the direction in Assumption 1.(iii) for the region B_i. Then*

$$\limsup_{\delta\to 0} I_{[0,T]}(\vec{r}_\delta) \leq I_{[0,T]}(\vec{r}). \tag{5.1}$$

Proof. This is proved in exactly the same way as Lemma 17. Let $\ell_1(t) = \ell(\vec{r}(t), \vec{r}'(t))$ and $\ell_2(t) = \ell(\vec{r}_\delta(t), \vec{r}'_\delta(t))$. Then, letting $\vec{\mu}(t)$ be the optimizing set of jump rates for the path $\vec{r}$, we have the equivalent of (4.18):

$$\ell_2(t) - \ell_1(t) \leq k \cdot K'_\lambda(\eta) + \sum_{i=1}^{k} \mu_i^*(t) \log \frac{\lambda_i^1(t)}{\lambda_i^2(t)}.$$

The same reasoning as in Lemma 17 then leads to a bound like (4.27), which immediately leads to the result. ∎

Lemma 27 *Assume 1–2. Using the notation therein, fix i, $T > 0$ and a path $\vec{r}(t)$ which takes values in B_i such that $I_{[0,T]}(\vec{r}) = K < \infty$. Then (with $\vec{r}(0) = \vec{x}$)*

$$\lim_{\delta\downarrow 0} \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}_x\left(\vec{z}^n(t) \in B(\vec{r}, \delta)\right) \geq -I_{[0,T]}(\vec{r}). \tag{5.2}$$

Proof. Let $\vec{v} = \vec{v}_i$ be the direction of the interior cone (see Assumptions 1–2) of B_i and let K_i be the constant such that for $\vec{x} \in B_i$ we have $d(\vec{x} + t\vec{v}, \partial G) \geq K_i t$ for t small. Denote by $\tilde{\eta}$ the modulus of continuity of $\vec{r}$ and set $\eta(a) = \max\{\tilde{\eta}(a), a\}$ so that $\eta^{-1}(a) \leq a$.
Now fix δ and set $t_\delta = \eta^{-1}(\delta/3)$. Then $t_\delta \leq \delta$ and for $t \leq t_\delta$,

$$\sup_{0\leq t\leq t_\delta} |t \cdot \vec{v} - \vec{r}(t)| \leq t_\delta \cdot |\vec{v}| + \eta(t_\delta) \leq \frac{2\delta}{3}. \tag{5.3}$$

Therefore, for $0 < \alpha < K_i/6$,

$$\mathbb{P}_x\left(\vec{z}^n \in B(\vec{r}, \delta)\right) \geq \mathbb{P}_x\left(|\vec{z}^n(t) - t \cdot \vec{v}| \leq \alpha\delta,\ 0 \leq t \leq t_\delta,\ \vec{z}^n \in B(\vec{r}, \delta)\right) \tag{5.4}$$

where the last ball is around the restriction of $\vec{r}$ to $[t_\delta, T]$. Now let $r_\delta(t) \triangleq \vec{r}(t) + t_\delta\vec{v}$ be a function on $[t_\delta, T]$. Then on this time interval

$$\sup_{t_\delta\leq t\leq T} |\vec{r}(t) - \vec{r}_\delta(t)| \leq \frac{\delta}{3}$$

and moreover, $d(\vec{r}_\delta(t), \partial G) \geq K_i\delta$. Therefore, for any function $\vec{u}$ on that time interval, $\|\vec{u} - \vec{r}_\delta\| \leq K_i\delta/2$ implies that $\|\vec{u} - \vec{r}\| \leq 2\delta/3$ and $d(\vec{u}(t), \partial G) \geq K_i\delta/2$. Now let $B_\delta \triangleq B(\vec{x} + t_\delta\vec{v}, \alpha\delta)$ and let $\vec{r}^{\,y}_\delta$ be the shift of $\vec{r}_\delta$ so that $\vec{r}^{\,y}_\delta(t_\delta) = \vec{y}$. Then

$$\mathbb{P}_x\left(\vec{z}^n \in B(\vec{r}, \delta)\right) \geq \mathbb{P}_x\left(|\vec{z}^n(t) - t\cdot\vec{v}| \leq \frac{\delta}{6},\ 0 \leq t \leq t_\delta\right) \times \inf_{y\in B_\delta} \mathbb{P}_y\left(\vec{z}^n \in B(\vec{r}^{\,y}_\delta, K_i\delta/2)\right) \tag{5.5}$$

where the last ball contains paths on $[t_\delta, T]$. Now the first term is bounded below by the probability that, over $0 \leq t \leq t_\delta$, rates for jumps in directions outside the cone of Assumption 2.C are zero while the rates for jump in directions within the cone are such that the process proceeds with speed one. However, the second condition satisfies a standard large deviations lower bound, and so

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}_x\left(|\vec{z}^n(t) - t\cdot\vec{v}| \leq \alpha\delta,\ 0 \leq t \leq t_\delta\right) \geq -Ct_\delta \tag{5.6}$$

for some constant C. Consider now the second probability in (5.5). Since the paths in $B(\vec{r}^{\,y}_\delta, K_i\delta/2)$ are bounded away from the boundary uniformly in $\vec{y} \in B_\delta$ we have by [7, Thm. 5.51] that the large deviations lower bound holds uniformly over $\vec{y} \in B_\delta$ (where uniformity is the usual sense of analysis). Therefore

$$\liminf_{n\to\infty} \frac{1}{n} \log \inf_{y\in B_\delta} \mathbb{P}_y\left(\vec{z}^n \in B(\vec{r}^{\,y}_\delta, K_i\delta/2)\right) \tag{5.7}$$

$$\geq -\inf_{y\in B_\delta} \inf\{I_{[t_\delta,T]}(\vec{w}) : \vec{w} \in B(\vec{r}^{\,y}_\delta, K_i\delta/2) \tag{5.8}$$

$$\geq -I_{[t_\delta,T]}(\vec{r}_\delta) \tag{5.9}$$

$$\geq -I_{[0,T]}(\vec{r}_\delta) \tag{5.10}$$

$$\geq -I_{[0,T]}(\vec{r}), \tag{5.11}$$

where the final inequality comes from Lemma 26. Thus the result follows from (5.6)–(5.7). ∎

Proof of the lower bound of Theorem 2. Using Lemma 27 the proof of the lower bound is almost identical to that of Lemma 18, and so we only give a sketch of the proof. Given a path $\vec{r}$ with $I_{[0,T]}(\vec{r}) = K < \infty$ we break the path into J segments and use the shift-and-stitch argument of Lemma 18. Lemma 27 shows that the shifted path provides a good approximation for the rate and a lower bound for the probabilities, and that the probabilities of the shifts are bounded below (on the exponential scale) by an arbitrarily small constant. This establishes the lower bound. ∎

6 Interior zeros

In this section we state and prove a large deviations principle for processes that have rates that may become zero in the interior of the region G, not at a boundary. Our main assumption about these processes is that the cone $\mathcal{C}_{\vec{x}}$ of jump directions of the process does not change, but remains $\mathbb{R}^d$ for all $\vec{x}$. This is a strong assumption. But the result is strong, too: under Lipschitz continuity of the jump rates $\lambda_i(\vec{x})$, the large deviations principle holds, with the usual rate function. This is, therefore, a strict generalization of the large deviations principle proved in [7, Chapter 5], where the assumption is the logarithms of the jump rates is bounded, which itself implies $\mathcal{C}_{\vec{x}} = \mathbb{R}^d$ for all $\vec{x}$. Combined with the results for rates that diminish towards a boundary, we obtain a fairly general theory of diminishing rates, Corollary 4.

We state our theorem for Lipschitz continuous jump rates and processes that have no boundaries. But, using the exponential tightness argument of Corollary 7 we give proofs only for bounded regions and bounded Lipschitz jump rates. Our proofs are based very heavily on the arguments in [7]. Rather than reproduce those arguments, we give only the lemmas and arguments needed to extend the previous proof to the present case.

We begin with some notation. For $\delta > 0$ we define

$$\lambda_i^\delta(\vec{x}) \triangleq \begin{cases} \lambda_i(\vec{x}) & \text{if } \lambda_i(\vec{x}) > \delta \\ 0 & \text{otherwise.} \end{cases} \tag{6.1}$$

We define $\mathcal{C}_{\vec{x}}^\delta$ as the positive cone spanned by the $\vec{e}_i$ whose corresponding $\lambda_i^\delta(\vec{x}) > 0$. Let $\mathcal{I}^\nu(\vec{x}) = \{i : \lambda_i(\vec{x}) > \nu\}$, so that $\mathcal{I}^0(\vec{x}) = \{i : \lambda_i(\vec{x}) > 0\}$. Also, for a set $\mathcal{I}$ of indices, we let $\mathcal{C}_{\vec{x}}(\mathcal{I})$ be the positive cone spanned by $\{\vec{e}_i : i \in \mathcal{I}, \lambda_i(\vec{x}) > 0\}$.

Lemma 28 *Assume 3 and that the $\lambda_i(\vec{x})$ are continuous. Then for every $R < \infty$, there exists $\delta > 0$ and $\nu > 0$ such that for $|\vec{x}| \le R$ we have a set $\mathcal{I}(\vec{x})$ with the following property. Every $\vec{z} \in B(\vec{x}, \delta)$ has $\lambda_i(\vec{z}) > \nu$, $i \in \mathcal{I}(\vec{x})$ and $\mathcal{C}_{\vec{z}}(\mathcal{I}(\vec{x})) = \mathbb{R}^d$.*

Note: we use this ν in following lemmas and proofs.

Proof. This follows easily from compactness. The set S, consisting of $\vec{x}$ such that at least one $\lambda_i(\vec{x}) = 0$, is closed. Therefore, the set of such $\vec{x}$ that satisfy $|\vec{x}| \leq R$ is compact. Cover each point $\vec{x}$ in S with an open ball with radius δ chosen so that the minimal rate of $\lambda_i(\vec{z}), i \in \mathcal{I}^0(\vec{x}),\ \vec{z} \in B(\vec{x}, \delta)$ is at least half the minimal rate of $\lambda_i(\vec{x}), i \in \mathcal{I}^0(\vec{x})$. By assumption, $\mathcal{C}_{\vec{x}} = \mathbb{R}^d$, so $\mathcal{C}_{\vec{z}}(\mathcal{I}^0(\vec{x})) = \mathbb{R}^d$ for all $\vec{z} \in B(\vec{x}, \delta)$. Choose a finite subcover of such balls $\{B(\vec{x}_j, \delta)\}$. Call the resulting union of these sets U. Then

$$\inf\left\{\lambda_i(\vec{z}) : \vec{z} \in B(\vec{x}_j, \delta_j),\ i \in \mathcal{I}^0(\vec{x}_j)\right\} \stackrel{\triangle}{=} \eta > 0$$

by construction. Then for $\vec{x} \notin U$, $\lambda_i(\vec{x}) > 0$. In fact, there is a positive bound, which without loss of generality we take to be η such that $\lambda_i(\vec{x}) \geq \eta$ for all $\vec{x} : |\vec{x}| \leq R, \vec{x} \notin U$, since this set is closed and the rates are continuous and not equal to 0. Now set $\nu = \eta/2$. The proof is concluded by showing that there is a δ so that each $\vec{x} \in U$,

$$B(\vec{x}, \delta) \cap U \subset B(\vec{x}_j, \delta_j) \text{ for some } j;\ \text{set } \mathcal{I}(\vec{x}) = \mathcal{I}(\vec{x}_j) \tag{6.2}$$

$$\text{or } d(\vec{x}, \partial U) \leq \delta;\ \text{set } \mathcal{I}(\vec{x}) = \{1, \dots, J\}. \tag{6.3}$$

The first case, equation (6.2), obviously satisfies the statement of the lemma. In the second case, (6.3), since the rates are continuous over a compact set, they are uniformly continuous, and we choose δ so that the rates change by at most $\eta/2$ over the δ ball. We establish the claim by contradiction; assume the contrary. Then there is a sequence $\vec{x}_i$ and $\delta_i \downarrow 0$ so that $\vec{x}_i \in U$ and the ball $B(\vec{x}_i, \delta_i)$ is not contained in any $B(\vec{x}_j, \delta_j)$. Take a converging subsequence with limit $\vec{x}$: then if $\vec{x} \in U$ we obtain a contradiction, since U is a finite union of balls, so that, for large i, the ball around $\vec{x}_i$ must be in some $B(\vec{x}_j, \delta_j)$. If $\vec{x}$ is on the boundary of U then for large i, $\vec{x}_i$ is within δ of ∂U. ∎

We define $\vec{\mu}^*(\vec{x}, \vec{y})$ as the unique optimizer for $f(\vec{\mu}, \vec{\lambda})$; that is, $\vec{\mu}$ causes $\ell(\vec{x}, \vec{y}) = f(\vec{\mu}, \vec{\lambda}(\vec{x}))$, with $\vec{y} = \sum_i \mu_i \vec{e}_i$. We define $\vec{\mu}^*_\delta(\vec{x}, \vec{y})$ to be the optimizer for rates λ^δ. We also define $\ell_\delta(\vec{x}, \vec{y})$ (not to be confused with ℓ^δ!) to be the rate function with jump rates λ^δ.

Lemma 29 *Assume 3 and that the λ_i are continuous. The maximizing $\vec{\theta}$ in the definition of ℓ is bounded uniformly in $|\vec{x}| \leq R$ and $|\vec{y}| \leq C$.*

Proof. The proof is similar to [7, Lemma 5.21]. By Lemma 28,

$$\max\{\langle \vec{\theta}, \vec{e}_i \rangle : i \in \mathcal{I}^\nu(\vec{x})\} \geq \alpha |\vec{\theta}|$$

for some $\alpha > 0$ that depends only on R. Let $\vec{\theta}_n$ be a maximizing sequence. Representing $\vec{y}$ as in Lemma 11, and using Lemma 28,

$$\ell(\vec{x},\vec{y}) = \lim_{n\to\infty} \langle\vec{\theta}_n,\vec{y}\rangle - \sum_{i=1}^{k} \lambda_i(\vec{x}) \left(e^{\langle\vec{\theta}_n,\vec{e}_i\rangle} - 1\right) \tag{6.4}$$

$$\leq \lim_{n\to\infty} |\vec{\theta}_n| C - \nu e^{\alpha|\vec{\theta}_n|} + \sum_{i=1}^{k} \lambda_i(\vec{x}). \tag{6.5}$$

However, the last sum is bounded and the function $ax + c - e^x$ diverges to $(-\infty)$ as $x \to \infty$. Since ℓ is non negative we conclude that $|\vec{\theta}_n|$ must be bounded for large n, where the bound depends only on R, C, ν and $\mathcal{I}^\nu$. ∎

Corollary 30 *Assume 3 and that the λ_i are continuous. For each B there exists a C such that for all $|\vec{x}| \leq B$ and all $|\vec{y}| \leq B$,*

$$\ell(\vec{x},\vec{y}) \leq C. \tag{6.6}$$

Proof. [7, Theorem 5.26, and exercise 5.30] show that the optimizing μ^* can be represented as $\mu_i^* = \lambda_i e^{\langle\vec{\theta}^*,\vec{e}_i\rangle}$, where $\vec{\theta}^*$ is the optimizing $\vec{\theta}$ in the definition of ℓ. Therefore, since Lemma 29 shows that $\vec{\theta}^*$ is bounded for bounded $\vec{y}$, we have μ_i^*/λ_i is bounded for bounded $\vec{y}$. Therefore, for optimizing μ^*, there is a constant u such that

$$\lambda_i - \mu_i^* + \mu_i^* \log\frac{\mu_i^*}{\lambda_i} \leq \lambda_i u. \tag{6.7}$$

∎

Lemma 31 *Assume 3 and that the $\lambda_i(\vec{x})$ are continuous. Then for each B and $\varepsilon > 0$ there exists a $\delta > 0$ such that for all $|\vec{y}| \leq B$, $\vec{x}_\delta \in B(\vec{x},\delta)$, $\vec{y}_\delta \in B(\vec{y},\delta)$ we have*

$$|\ell(\vec{x},\vec{y}) - \ell_\delta(\vec{x}_\delta,\vec{y}_\delta)| < \varepsilon.$$

Proof. By Lemma 10,

$$\ell(\vec{x},\vec{y}) \triangleq \inf_{\mu\in K_y}\left(\sum_{i=1}^{k} \lambda_i(\vec{x}) - \mu_i + \mu_i \log\frac{\mu_i}{\lambda_i(\vec{x})}\right) \tag{6.8}$$

$$\leq \inf_{\mu\in K_y}\left(\sum_{i:\lambda_i^\delta>0} \lambda_i^\delta(\vec{x}) - \mu_i + \mu_i \log\frac{\mu_i}{\lambda_i^\delta(\vec{x})}\right) + \sum_{i:\lambda_i^\delta=0} \lambda_i \tag{6.9}$$

$$\leq \ell_\delta(\vec{x},\vec{y}) + k\delta \tag{6.10}$$

where the first inequality holds since the infimum is taken over a smaller set—where $\lambda_i^\delta = 0$ implies $\mu_i = 0$. We claim that to conclude the proof it suffices to establish that

$$\ell_\delta(\vec{x}_\delta, \vec{y}_\delta) \leq \ell(\vec{x}, \vec{y}) + \varepsilon, \tag{6.11}$$

since using (6.8) and interchanging the roles of $\vec{x}, \vec{y}$ with $\vec{x}_\delta, \vec{y}_\delta$ in (6.11) and then using (6.8) again we get

$$\begin{aligned} \ell_\delta(\vec{x}_\delta, \vec{y}_\delta) &\geq \ell(\vec{x}_\delta, \vec{y}_\delta) - k\delta & (6.12)\\ &\geq \ell_\delta(\vec{x}, \vec{y}) - k\delta - \varepsilon & (6.13)\\ &\geq \ell(\vec{x}_\delta, \vec{y}_\delta) - 2k\delta - \varepsilon\,. & (6.14) \end{aligned}$$

The difficulty in demonstrating (6.11) is that the $\vec{\mu}_\delta^*$ might be very different from the $\vec{\mu}^*$. But we can bound this difference, using the property that $\mathcal{C}_{\vec{x}} = \mathbb{R}^d$ for all $\vec{x}$. Let $\mathcal{I}^\nu(\vec{x})$ be the set of i with $\lambda_i(\vec{x}_\delta) \geq \nu$ for all $\vec{x}_\delta$ near $\vec{x}$; see Lemma 28. Note that for each δ there is a uniformly bounded vector $\vec{\alpha}$ with $\alpha_i = 0$ for $i \notin \mathcal{I}^\nu(\vec{x})$, such that

$$\sum_i (\mu_i^* + \delta\alpha_i)\, \vec{e}_i = \vec{y}_\delta. \tag{6.15}$$

There may be some nonzero components of the μ_i for $i \notin \mathcal{I}^\nu(\vec{x})$. However, by Lemma 12, since $\vec{y}_\delta$ is bounded, for any $\eta > 0$ we may choose δ small enough that if any of the $\lambda_i^\delta = 0$ then by continuity, $\lambda_i < \eta$. Take

$$\vec{v}_\delta = \sum_{i \notin \mathcal{I}} \mu_i^* \vec{e}_i. \tag{6.16}$$

Then $\vec{v}_\delta \in \mathcal{C}_{\vec{z}}^\delta$ for all $\vec{z} \in B(\vec{x}, \delta)$, so by Lemma 11 we may write

$$\vec{v}_\delta = \sum_{i \in \mathcal{I}} a_i \vec{e}_i \tag{6.17}$$

for some $a_i \geq 0$, $|a_i| \leq \kappa\delta$. Therefore we may write

$$\vec{y}_\delta = \sum_{i \in \mathcal{I}} (\mu_i^* + a_i + \delta\alpha_i)\, \vec{e}_i. \tag{6.18}$$

We have

$$\begin{aligned} &\ell(\vec{x}_\delta, \vec{y}_\delta) \\ &\quad \leq \sum_{i \in \mathcal{I}} \lambda_i(\vec{x}_\delta) - \mu_i^* - a_i - \delta\alpha_i + (\mu_i^* + a_i + \delta\alpha_i) \log \frac{\mu_i^* + a_i + \delta\alpha_i}{\lambda_i(\vec{x}_\delta)}. \end{aligned} \tag{6.19}$$

Note that

$$\ell(\vec{x}, \vec{y}) = \sum_i \lambda_i(\vec{x}) - \mu_i^* + \mu_i^* \log \frac{\mu_i^*}{\lambda_i}. \tag{6.20}$$

Both $\vec{\alpha}$ and $\vec{a}$ are of size δ, so the difference between corresponding terms in the sums (6.19) and (6.20) can be bounded by terms that go to zero with δ. Furthermore, by definition of $\mathcal{I}$, $\lambda_i(\vec{z}) \geq \nu$ for all $\vec{z}$ near $\vec{x}$. This finishes the estimates. ■

Proof of Theorem 3. The large deviations upper bound follows from these arguments based on the argument in [7]. The key theorem there is [7, Theorem 5.64]. It is based on Lemmas 5.57 and 5.58; Lemma 5.63; Lemma 5.62; and Lemma 5.48 there. Lemmas 5.57 and 5.58 require only bounded rates λ_i, and Lemma 5.63 requires bounded continuous rates, so these three lemmas continue to hold. Lemma 5.62 is based on Lemma 5.43, which required log-bounded rates, but is established without the lower bound on the rates in our Lemma 23 which requires only absolutely continuous jump rates λ_i. Finally, [7, Lemma 5.48] is based on Lemma 5.35 there, which also requires log-bounded rates. Our new Lemma 31 replaces Lemma 5.35. So, under the assumption that the $\lambda_i(\vec{x})$ are bounded, absolutely continuous, and that $\mathcal{C}_{\vec{x}} = \mathbb{R}^d$ for all $\vec{x}$, the large deviations upper bound is proved.

The large deviations lower bound follows even more directly. Assuming that the rates $\lambda_i(\vec{x})$ are bounded and Lipschitz continuous, so that Kurtz's theorem [7, Theorem 5.3] applies, essentially the same proof of the lower bound [7, Theorem 5.51], goes through. The only place that log-boundedness is used is [7, Corollary 5.53], and using Lemma 31 it is easily seen to hold without that assumption. Using Lemma 31, we see that the rate function $\ell(\vec{x}, \vec{y})$ is jointly continuous in $\vec{x}$ and $\vec{y}$ over bounded regions. Approximating jump rates with λ_i^δ makes a small (δ) difference in the rate function; this can be made arbitrarily small. So, under the assumption that the $\lambda_i(\vec{x})$ are bounded, Lipschitz continuous, and that $\mathcal{C}_{\vec{x}} = \mathbb{R}^d$ for all $\vec{x}$, the large deviations lower bound is proved. ■

7 Reachability of the boundary

We now provide a simple condition for showing that a point $\vec{x} \in \partial G$ may be reached with a finite cost (hence exponentially non-zero probability) via a path from the interior of G. We state a sufficient condition that is far from necessary; nevertheless, it is general enough to cover many cases of interest. For the one-dimensional case, we prove that a similar condition is also necessary. Recall that $s(\delta)$ is the scale function defined in Lemma 9.

Lemma 32 *Assume that the rates λ_i are bounded and let Assumption 2.C hold. Fix $\vec{x} \in \partial G$. If*

$$\int_0^\delta \log \frac{1}{s(t)}\, dt < \infty \tag{7.1}$$

for some $\delta > 0$ then $I_{[0,T]}(\vec{r}\,) < \infty$ for some $\vec{r}$ with $\vec{r}\,(T) = \vec{x}$.

Proof. For convenience we shift time. Let $\vec{r}\,(t)$ be a path with the following properties: $\vec{r}\,(0) = \vec{x}$; for some $c > 0$ and $t_0 < 0$ we have $d(\vec{r}\,(t), \partial G) > c|t|$ for $t_0 \le t \le 0$; and, for some $C > 0$, $\left|\vec{r}\,'(t)\right| < C$ for almost all $t \in [t_0, 0)$. Since $\left|\vec{r}\,'(t)\right| < C$ for almost all $t \in [t_0, 0)$, by Lemma 12 $\mu_i \le C_0$, where $\mu_i(t)$ is the optimal change of measure for $\ell(\vec{r}\,(t), \vec{r}\,'(t))$. Since we assume $d(\vec{r}\,(t), \partial G) > c|t|$ for $t_0 \le t \le 0$, we have from (3.1)

$$\ell(\vec{r}\,(t), \vec{r}\,'(t)) \le C + C \log \frac{1}{s(d(\vec{r}\,(t), \partial G))}\ , \tag{7.2}$$

since both λ_i and μ_i are bounded. But $d(\vec{r}\,(t), \partial G) > c|t|$. Therefore,

$$\ell(\vec{r}\,(t), \vec{r}\,'(t)) \le C + C \log \frac{1}{s(ct))} \tag{7.3}$$

This proves the result. ∎

Note that the condition for reachability holds for any polynomially decreasing scale function, since $\int_0^1 \frac{1}{\log x^j}\, dx < \infty$. In particular, in the case of an infinite server queue, where the rates decrease linearly to zero at the boundary, the boundary may be reached in finite time at finite cost. Note also that the condition is tight under the assumption that $\left|\vec{r}\,'(t)\right| < C$: the linear rate of approaching the boundary is optimal under this bound, as can be seen by a change of variable argument.

In the case of one-dimensional processes, the condition on reachability is virtually necessary as well as sufficient. This result is interesting enough that we detail it here. We suppose without loss of generality that the boundary is $x = 0$, and that the interior of G is contained in $x > 0$. Recall our assumption that at least one jump rate away from the boundary is strictly positive at the boundary. Let $\mu(x)$ denote the sum of the jump rates in negative directions, and $\Lambda(x)$ denote the sum in positive directions. We consider a path $r(t) = b - at$ for $t \in [0, T = b/a]$, where we suppose that $[0, b] \in G$.

Lemma 33 *Consider the one dimensional case. Let Assumption 2.C hold. Assume that the rates λ_i are bounded and that at least one rate of jump away from the boundary is bounded away from 0. Then the boundary can be reached by the path $r(t) = b - at$ with finite cost under the following condition:*

$$I_{[0,T]}(r) < \infty \textit{ if and only if } \int_0^b \log \frac{1}{\mu(x)}\, dx < \infty. \tag{7.4}$$

Proof.

$$\ell(r, r') = \sup_\theta \left(-a\theta - \sum_i \lambda_i(r)(\exp(\theta(-a)e_i) - 1) \right). \tag{7.5}$$

Let $\mathcal{I}^+$ be the i with $e_i > 0$, and let $\mathcal{I}^-$ be the i with $e_i < 0$. Differentiating (7.5) with respect to θ and setting the result equal to zero, we see

$$-a - \sum_{i\in\mathcal{I}^+} \lambda_i(r)e_i e^{\theta e_i} - \sum_{i\in\mathcal{I}^-} \lambda_i(r)e_i e^{\theta e_i} = 0, \tag{7.6}$$

or, equivalently,

$$\sum_{i\in\mathcal{I}^-} \lambda_i(r)(-e_i)e^{\theta e_i} = a + \sum_{i\in\mathcal{I}^+} \lambda_i(r)e_i e^{\theta e_i} \tag{7.7}$$

Both sides of (7.7) are positive, consisting of all positive terms. As $r \to 0$ we have $\lambda_i(r) \to 0$ for all $i \in \mathcal{I}^-$. Therefore we have $\theta \to -\infty$ as $r \to 0$; recalling that for at least one $j \in \mathcal{I}^+$ we have $\lambda_j(0) > 0$, we see that the rate at which $\theta \to -\infty$ as $r \to 0$ is bounded below, independent of a.

Let

$$h \triangleq \min_{i\in\mathcal{I}^-} |e_i| \text{ and } H \triangleq \max_{i\in\mathcal{I}^-} |e_i| \tag{7.8}$$

$$\varepsilon(r) \triangleq \sum_{i\in\mathcal{I}^+} \lambda_j(r)e_i e^{\theta e_i}. \tag{7.9}$$

Then $\varepsilon(r) \to 0$ as $r \to 0$, since $\theta \to -\infty$ as $r \to 0$. By (7.7) we have

$$\mu(r)he^{\theta h} \le a + \varepsilon(r) \le \mu(r)He^{\theta H} \tag{7.10}$$

(recall $\mu(r) = \sum_{i\in\mathcal{I}^-} \lambda_i(r)$). Therefore, for each $\delta > 0$, when r is close enough to zero, by using (7.10) to bound θ, we obtain

$$\ell(r, r') \le \frac{a}{h} \log \frac{a + \varepsilon(r)}{\mu h} + \mu + \Lambda + \delta \tag{7.11}$$

$$\ell(r, r') \ge \frac{a}{H} \log \frac{a + \varepsilon(r)}{\mu H} + \mu + \Lambda - \delta - \frac{a + \varepsilon(r)}{h}; \tag{7.12}$$

the last term on the right of (7.12) is derived from (7.7) by the estimate

$$\sum_i \lambda_i(r)e^{\theta e_i} \leq \frac{a+\varepsilon(r)}{h}.$$

Therefore

$$I_{[0,T]}(r) = \int_0^T \left(\log \frac{-a}{\mu(r(t))} + O(1)\right) dt \tag{7.13}$$

$$= \int_0^b \left(\log \frac{1}{\mu(r)} + O(1)\right) dt. \tag{7.14}$$

This shows that $I_{[0,T]}(r)$ is finite if and only if (7.14) is. ∎

8 Conclusion

We have shown that the usual sample-path jump Markov large deviations theorem and rate function remain unchanged even when jump rates tend to zero in some cases. The cases include the very important one of infinite server queues, as well as the case where the positive cone of jump directions does not change. Moreover, while existing theories that include boundaries assume flat boundaries, the boundaries here are quite general.

However, our understanding of diminishing rate is not complete, as illustrated in Example C of the Introduction. Moreover, a theory combining diminishing rates with discontinuous rates, even with flat boundaries, or finite level boundaries, is still lacking.

9 Appendix: Extensions to non-convex sets

In Section 2 we proved a "uniform cone condition" for convex sets (Lemma 1), and showed how to extend the rates to $\mathbb{R}^d$ when G is convex. In this section we extend these results to more general sets, of the type that covers most applications.

Let G be a closed set which is the union of a finite number of convex sets G_j. Assume that any finite intersection $\cap_{j\in J_\alpha} G_j$ is either empty or is the closure of its interior. It can then be shown that Assumption 1 holds. This is the case since any non-empty intersection of G_j is convex, and so we can apply the argument for convex sets. This together with a continuity property yields this result.

Note that the intersection condition precludes the presence of external cusps.

In the rest of this section we show that the second part of Assumption 2.A follows from the first for unions of convex sets. We use the following notation. G_j are closed convex sets, and for any point x denote by $y_j(x)$ the projection of x into the (closure of) G_j.

Lemma 34 (Contraction) *For any x, z and j,*

$$|y_j(x) - y_j(z)| \le |x - z| \,. \tag{9.1}$$

Proof. Obvious. But can be made precise as follows. The only non trivial case is where both are outside G_j. But then

$$|x - y_j(z)|^2 > |x - y_j(x)|^2 + |y_j(x) - y_j(z)|^2 \tag{9.2}$$

since the angle $xy_j(x)y_j(z)$ is necessarily larger than 90 degrees. Write the symmetric expression and sum up to obtain the result. ∎

Lemma 35 (Continuity) *Let f be a Lipschitz function on G with Lipschitz constant L_f. Define the function g through*

$$g(x) = \begin{cases} f(x) & \text{for } x \in G\,, \\ \dfrac{\sum_{j=1}^{J} (d[x, y_j(x)])^{-1} f(y_j(x))}{\sum_{j=1}^{J} (d[x, y_j(x)])^{-1}} & \text{for } x \notin G\,. \end{cases} \tag{9.3}$$

Then g is continuous.

Proof. Fix an arbitrary point x and note that it suffices to prove continuity locally at x. This is obvious if x is in the interior of G or in the interior of its complement. Since G is closed, the only case to consider is how g changes between the boundary and an outside point. So let $\{x_n\}$ be a sequence of points outside G converging to $x \in G$. Suppose first that $x \in G_j$ but $x \notin G_k,\ k \neq j$. Then by lemma 34 $y_j(x_n) \to y_j(x) = x$. Since G_k is closed, we have $y_k(x_n) > \varepsilon_k \ge \varepsilon > 0$ for some ε, all $k \neq j$ and all n large. Continuity then follows since $\{f(x_n)\}$ is bounded. In the general case, by reordering the indices we may assume that, for some ℓ, $x \in G_j$ for all $j \le \ell$ and $x \notin G_j$ for all $j > \ell$. Since f is Lipschitz and $y_j(x) = x$, lemma 34 gives

$$f(y_j(x_n)) = f(x) + \varepsilon_{jn}\,, \quad |\varepsilon_{jn}| \le L_f |x_n - x| \quad \text{for } j \le \ell, \tag{9.4}$$

while obviously

$$|f(y_j(x_n)) - f(x)| \leq K \quad \text{for some } K\text{, for } j > \ell. \tag{9.5}$$

By definition, for some $\varepsilon > 0$ we have, for all large n,

$$\begin{cases} d(x_n, y_j(x_n)) \to 0 & j \leq \ell\text{, and} \\ d(x_n, y_j(x_n)) > \varepsilon & j > \ell. \end{cases} \tag{9.6}$$

Let $d_n \stackrel{\triangle}{=} \min_j\{d(x_n, y_j(x_n))\}$. Then $0 < d_n(d[x_n, y_j(x_n)])^{-1} \leq 1$, and

$$\lim_{n\to\infty} d_n(d[x_n, y_j(x_n)])^{-1}\varepsilon_{jn} = 0 \quad \text{for } j \leq \ell, \tag{9.7}$$

$$\lim_{n\to\infty} d_n(d[x_n, y_j(x_n)])^{-1}K = 0 \quad \text{for } j > \ell. \tag{9.8}$$

$$\tag{9.9}$$

Therefore

$$\lim_n g(x_n) = \lim_n \frac{\sum_{j=1}^J d_n(d[x_n, y_j(x_n)])^{-1} f(y_j(x_n))}{\sum_{j=1}^J d_n(d[x_n, y_j(x_n)])^{-1}} \tag{9.10}$$

$$= \lim_n \frac{\sum_{j=1}^\ell d_n(d[x_n, y_j(x_n)])^{-1} f(x)}{\sum_{j=1}^\ell d_n(d[x_n, y_j(x_n)])^{-1}} \tag{9.11}$$

$$= f(x)\,. \tag{9.12}$$

Thus $g(x)$ is continuous at each point x. ∎

We now turn to a proof of Lipschitz continuity. This, again, is a local property.

Lemma 36 (Lipschitz) *Under the conditions of Lemma 35, g as defined in* (9.3) *is Lipschitz continuous with Lipschitz constant L_f.*

Proof. It suffices to prove Lipschitz continuity with constant L_f locally at x for each point x. This holds by definition for x in the interior of G, and we next establish the result for x in the interior of its complement. So, fix such a point x and define

$$\varepsilon \stackrel{\triangle}{=} \frac{1}{2}\min\{d(x,z) : z \in G\} \tag{9.13}$$

so that by our assumption $\varepsilon > 0$. Fix an arbitrary $0 < \delta < 1/2$ and a point z such that $d(z,x) < \delta\varepsilon$: then $d(z,G) > \varepsilon$. Since x is fixed it will be convenient to denote $q_j \triangleq (d[x, y_j(x)])^{-1}$. Without loss of generality we assume that $f(x), f(y_j(x)), f(z)$ and $f(y_j(z))$ are all positive (this amounts to a shift by a constant, and does not influence continuity properties). Now by lemma 34,

$$\begin{aligned} d[x, y_j(x)] &\leq d[x, z] + d[z, y_j(z)] + d[y_j(z), y_j(x)] && (9.14) \\ &\leq d[z, y_j(z)] + 2\varepsilon\delta && (9.15) \end{aligned}$$

so that, since $d(x, G) > \varepsilon$,

$$\begin{aligned} \frac{1}{d[z, y_j(z)]} &\leq \frac{1}{d[x, y_j(x)] - 2\varepsilon\delta} && (9.16) \\ &\leq q_j(1 + 2\delta)\,. && (9.17) \end{aligned}$$

Exchanging the roles in the triangle inequality, we conclude that

$$q_j(1 - 2\delta) \leq \frac{1}{d[z, y_j(z)]} \leq q_j(1 + 2\delta)\,. \qquad (9.18)$$

But then

$$\begin{aligned} g(x) - g(z) &= \frac{\sum_{j=1}^{J} q_j f(y_j(x))}{\sum_{j=1}^{J} q_j} - \frac{\sum_{j=1}^{J} (d[z, y_j(z)])^{-1} f(y_j(z))}{\sum_{j=1}^{J} (d[z, y_j(z)])^{-1}} && (9.19) \\ &\leq \frac{\sum_{j=1}^{J} q_j f(y_j(x))}{\sum_{j=1}^{J} q_j} - \frac{\sum_{j=1}^{J} q_j (1 - 2\delta) f(y_j(z))}{\sum_{j=1}^{J} q_j (1 + 2\delta)} && (9.20) \\ &\leq \frac{\sum_{j=1}^{J} q_j [f(y_j(x)) - f(y_j(z))]}{\sum_{j=1}^{J} q_j} + \frac{\sum_{j=1}^{J} q_j f(y_j(z))}{\sum_{j=1}^{J} q_j} \left(1 - \frac{1 - 2\delta}{1 + 2\delta}\right) && (9.21) \\ &\leq L_f d(x, z) + \max_j \{|f(y_j(z))|\} \cdot \frac{4\delta}{1 + 2\delta}\,. && (9.22) \end{aligned}$$

Exchanging the roles of x and z we obtain in exactly the same way

$$g(z) - g(x) \leq L_f d(x, z) + \max_j \{|f(y_j(x))|\} \cdot \frac{4\delta}{1 + 2\delta}\,. \qquad (9.23)$$

Since δ was arbitrary, the Lipschitz continuity is established.

It remains to consider the case where x is on the boundary of G. So fix a direction z: if $x + \eta \cdot z$ is in G for all η small, then there is nothing to prove. We need only consider the case where $x + \eta \cdot z$ is in the (open) complement of G for all $0 < \eta \leq \eta_0$. So fix an arbitrary ε. By lemma 35, we can find a δ and a point u in the complement of G (actually, on the line segment $(x, x + \eta_0 z)$) so that

$$|g(x) - g(u)| \leq \varepsilon\,, \tag{9.24}$$

$$d(x, u) \leq \varepsilon d(x, z)\,. \tag{9.25}$$

By the Lipschitz property for points in the complement of G we conclude

$$|g(x) - g(z)| \leq |g(x) - g(u)| + |g(u) - g(z)| \tag{9.26}$$

$$\leq \varepsilon + L_f d(u, z) \tag{9.27}$$

$$\leq \varepsilon + L_f(1 + \varepsilon) d(x, z)\,. \tag{9.28}$$

Since ε is arbitrary the result is established. ■

References

[1] Atar, Rami and Paul Dupuis, "Large deviations and queueing networks: methods for rate function identification," *Stoch. Proc. Appl.* **84** pp. 255–296, 1999.

[2] Dupuis, Paul and Richard Ellis, "The large deviations principle for a general class of queueing systems I." *Trans. Amer. Math. Soc.* **347** pp. 2689–2751, 1995.

[3] Dupuis, Paul, Richard Ellis and Alan Weiss, "Large deviations for Markov processes with discontinuous statistics I: general upper bounds," *Ann. Prob.* **19** pp. 1280–1297, 1991.

[4] Dembo, Amir and Ofer Zeitouni, *Large Deviations Techniques and Applications*, Second Edition, Springer-Verlag, NY 1998.

[5] Ignatiouk-Robert, Irina, "Sample path large deviations and convergence parameters," *Ann. Applied Prob.* **11** pp. 1292–2329, 2002.

[6] Mandjes, Michel and Alan Weiss, "Sample path large deviations of a multiple time-scale queueing model," submitted, 2000.

[7] Shwartz, Adam and Alan Weiss, *Large Deviations for Performance Analysis*, Chapman-Hall, 1995.

[8] Tse, David, "Asymptotic Optimality of a Measurement-Based Call Admission Control Scheme", Proc. 34th Allerton Conf, Oct., 1996.