The Annals of Applied Probability
1998, Vol. 8, No. 1, 45–66

ON LARGE DEVIATIONS OF MARKOV PROCESSES WITH DISCONTINUOUS STATISTICS[1]

BY MURAT ALANYALI[2] AND BRUCE HAJEK

Bell Laboratories and University of Illinois, Urbana–Champaign

This paper establishes a process-level large deviations principle for Markov processes in the Euclidean space with a discontinuity in the transition mechanism along a hyperplane. The transition mechanism of the process is assumed to be continuous on one closed half-space and also continuous on the complementary open half-space. Similar results were recently obtained for discrete time processes by Dupuis and Ellis and by Nagot. Our proof relies on the work of Blinovskii and Dobrushin, which in turn is based on an earlier work of Dupuis and Ellis.

1. Overview of previous work. Large deviations of Markov processes with discontinuous transition mechanisms arise in a broad range of applications such as the analysis of load sharing and queueing networks [see Alanyali and Hajek (1998) and Dupuis and Ellis (1995) for examples]. This paper establishes a large deviations principle (LDP) for a Markov random process in R^d with a discontinuity in the transition mechanism along a hyperplane. The transition mechanism of the process is assumed to be continuous on one closed half-space, and also continuous on the complementary open half-space. The following paragraphs give an overview of the related work and identify the contribution of the present paper. The formulation and proof of the main result are the subjects of subsequent sections.

In their paper Dupuis and Ellis (1992) established an explicit representation of the rate function in the case of constant transition mechanism in the two half-spaces. The paper proved an LDP for the process observed at a fixed point in time, though an underlying process-level LDP is implicit in the paper.

Subsequently, Blinovskii and Dobrushin (1994) and Ignatyuk, Malyshev, and Scherbakov (1994) derived process-level LDP's for the case of constant transition mechanism in each half-space, using different approaches. The work by Ignatyuk, Malyshev, and Scherbakov (1994) is somewhat restrictive in that the first coordinate of the process is assumed to take values in a lattice, and, when off the hyperplane, the process can step at most one unit towards the hyperplane at a time. This condition prevents jumps that strictly cross the hyperplane of discontinuity. On the other hand, the work allows the process to have a different transition mechanism in each open half-space and on the hyperplane itself. The paper by Blinovskii and Dobrushin (1994) does not rely

Received January 1996; revised May 1997.

[1]Research partially supported by the NSF NCR 93-14253.

[2]Completed while a student at the University of Illinois with partial support by a TUBITAK NATO Fellowship.

AMS 1991 *subject classifications*. Primary 60F10; secondary 60J.

Key words and phrases. Large deviations, Markov processes, discontinuous statistics.

on a lattice assumption—the jump distribution need not even be concentrated on a countable number of points, and jumps strictly crossing the boundary can occur.

Shwartz and Weiss (1995) established an LDP for a process on a half-space with a flat boundary which cannot be crossed. The transition mechanism can vary continuously on both the open half-space and on the hyperplane boundary. The method is lattice based and also includes the assumption of at most unit jumps towards the boundary. The model applies to processes with continuous transition mechanisms in two half-spaces separated by a hyperplane only if a symmetry condition holds. A somewhat different explicit representation for the rate function is given in Shwartz and Weiss (1995), though as noted by Remark 2.2 below, it can be easily related to the expression of Dupuis and Ellis (1992). Dupuis, Ellis, and Weiss (1991) established a large deviations style upper bound, which is tight for the flat boundary process of Shwartz and Weiss (1995), but which for the case of two half-spaces separated by a boundary is not always tight.

Dupuis and Ellis (1995) established LDP's for Markov processes with transition probabilities that are continuous over facets generated by a finite number of hyperplanes. For example, two intersecting hyperplanes generate nine such facets. While in general the paper does not identify the rate function explicitly, it does state an explicit integral representation for the case of a single hyperplane of discontinuity. The integrand in the representation given by Dupuis and Ellis (1995) and Blinovskii and Dobrushin (1994) has the form established in Dupuis and Ellis (1992). It is assumed in Dupuis and Ellis (1995) that the processes are lattice valued and satisfy a mild communication–controllability condition.

The LDP established in this paper is based on an adaptation of the construction in Blinovskii and Dobrushin (1994), hereafter referred to as BD. As with BD, it therefore does not require lattice assumptions. The method described in Dupuis and Ellis (1995) accommodates the continuous variation of transition mechanisms throughout the proof, while it is not clear how to directly incorporate continuous variation of transition mechanisms in the approach of BD. The tack taken in this paper, therefore, is a two-step procedure: first an LDP for a piecewise-constant transition mechanism is identified, and then the LDP is extended to cover a continuously varying transition mechanism within the half-spaces.

Another contribution of this paper is to somewhat streamline BD's proof and to show that the method is appropriate in either continuous or discrete time.

Since the original submission of this paper, a reviewer pointed out to us the existence of the work of Nagot (1995), and the book of Dupuis and Ellis (1996) appeared. Both of these works provide large deviations results for discrete time Markov processes with statistics varying continuously on each side of a hyperplane. The restriction on the degree of variability of the statistics within each half-space imposed in both these papers is continuity in the topology of weak convergence of probability measures, which is more general than the uniform continuity condition that we require. Nagot (1995) requires each jump measure to have bounded support, among other assumptions. The only assumption of Dupuis and Ellis (1996) not required in our Theorem 2.2 is a mild technical assumption that the relative interior of the closed convex hull of the support of the jump measures (such convex hull is assumed to be constant in each half-space) contain the zero vector. Nagot (1995) provides comparisons of work on large deviations for processes with continuously varying statistics, starting with the pioneering work of Azencott and Ruget (1977).

To date, none of the published works explicitly treat processes with discontinuities along a *curved* separating surface. In this connection, the formulation of Azencott and Ruget (1977) is interesting in that it is appropriate for processes (with continuous statistics) on differential manifolds. In particular, the formulation of Azencott and Ruget (1977) allows the statistics of the *directions* of jumps from a starting point x to depend on the large parameter γ (but converge as $\gamma \to \infty$) as opposed to being constant. This is required so that the class of models is preserved under smooth changes of coordinates. Adoption of the Azencott and Ruget (1977) formulation for processes with discontinuous statistics would provide a natural formulation of processes with discontinuities along a curved surface.

2. Statement of the main result. Given a positive integer d, let R^d denote the d-dimensional Euclidean space. A collection $\nu = (\nu(x)\colon x \in R^d)$ is called a *rate-measure field* if for each x, $\nu(x) = \nu(x, \cdot)$ is a positive Borel measure on R^d and $\sup_x \nu(x, R^d) < \infty$. For each positive scalar γ, a right continuous Markov jump process $X^\gamma = (X_t^\gamma\colon t \geq 0)$ is said to be *generated* by the pair (γ, ν) if given its value at time t, the process X^γ jumps after a random time exponentially distributed with parameter $\gamma\nu(X_t^\gamma, R^d)$, and the jump size is a random variable Δ where $\gamma\Delta$ has distribution $\nu(X_t^\gamma)/\nu(X_t^\gamma, R^d)$, independent of the past history. The *polygonal interpolation* of the process X^γ, $\tilde{X}^\gamma$, is defined as

$$\tilde{X}_t^\gamma = \frac{t - \tau_k}{\tau_{k+1} - \tau_k} X_{\tau_{k+1}}^\gamma + \frac{\tau_{k+1} - t}{\tau_{k+1} - \tau_k} X_{\tau_k}^\gamma, \qquad \tau_k \leq t \leq \tau_{k+1},$$

where τ_k is the kth jump time of X^γ. Since X^γ has a finite number of jumps in bounded time intervals, $\tilde{X}^\gamma$ has sample paths in $C_{[0,\infty)}(R^d)$, the space of continuous functions $\phi\colon [0, \infty) \to R^d$ with the topology of uniform convergence on compact sets.

The following are some standard definitions of large deviations theory. Let $\mathscr{X}$ be a topological space and let Z^γ denote an $\mathscr{X}$-valued random variable for each $\gamma > 0$. The sequence $(Z^\gamma\colon \gamma > 0)$ is said to satisfy the large deviations principle with rate function $\Gamma\colon \mathscr{X} \to R_+ \cup \{\infty\}$ if Γ is lower semicontinuous, and for any Borel measurable $S \subset \mathscr{X}$,

$$\limsup_{\gamma \to \infty} \gamma^{-1} \log P(Z^\gamma \in S) \leq -\inf_{z \in \bar{S}} \Gamma(z),$$

$$\liminf_{\gamma \to \infty} \gamma^{-1} \log P(Z^\gamma \in S) \geq -\inf_{z \in S^o} \Gamma(z),$$

where $\overline{S}$ and S^o denote respectively the closure and the interior of S. The rate function Γ is called *good* if for each $l \geq 0$ the level set $\{z: \Gamma(z) \leq l\}$ is compact.

Let A^o denote the hyperplane $\{x \in R^d: x(1) = 0\}$, and set $A^+ = \{x \in R^d: x(1) > 0\}$ and $A^- = \{x \in R^d: x(1) < 0\}$. Given two rate-measure fields ν^+ and ν^-, let Λ^+, Λ^-, and Λ^o be defined as follows:

$$
\begin{aligned}
M^{\pm}(x, \zeta) &= \int_{R^d} (e^{z\zeta} - 1)\nu^{\pm}(x, dz), \qquad x, \zeta \in R^d, \\
\Lambda^{\pm}(x, y) &= \sup_{\zeta \in R^d} \{y\zeta - M^{\pm}(x, \zeta)\}, \qquad y \in R^d,
\end{aligned}
\tag{2.1}
$$

$$
\Lambda^o(x, y) = \inf_{\substack{0 \leq \beta \leq 1,\ y^+ \in \overline{A^-},\ y^- \in \overline{A^+} \\ \beta y^+ + (1-\beta) y^- = y}} \{\beta \Lambda^+(x, y^+) + (1-\beta)\Lambda^-(x, y^-)\}.
\tag{2.2}
$$

Consider the following conditions regarding rate-measure fields.

CONDITION 2.1 (Boundedness). There exists a finite number m such that $\nu(x, R^d) \leq m$ for all $x \in R^d$.

CONDITION 2.2 (Exponential moments). For each $\zeta \in R^d$ there exists a finite number b such that $\int_{R^d} (e^{z\zeta} - 1)\nu(x, dz)/\nu(x, R^d) < b$ for all $x \in R^d$.

CONDITION 2.3 (Uniform continuity). For each $x, x' \in R^d$ the measures $\nu(x)$ and $\nu(x')$ are equivalent. Furthermore, given a positive number ε, there exists a corresponding positive number δ such that $(1+\varepsilon)^{-1} \leq d\nu(x)/d\nu(x') \leq (1+\varepsilon)$ whenever $|x - x'| < \delta$.

The main result of the paper is the following theorem:

THEOREM 2.1. *Let ν^+ and ν^- be two rate-measure fields on R^d each of which satisfies Conditions* 2.1–2.3, *and $\nu^+(x, A^-) > 0$ and $\nu^-(x, A^+) > 0$ for some (equivalently all) $x \in R^d$. Let X^γ denote the Markov process generated by the pair (γ, ν), where ν is given by*

$$
\nu(x) = \begin{cases} \nu^+(x), & \text{if } x \in \overline{A^+}, \\ \nu^-(x), & \text{if } x \in A^-, \end{cases}
$$

and let $\tilde{X}^\gamma$ denote the polygonal interpolation of X^γ. Suppose $\tilde{X}_0^\gamma = x^\gamma$ where $(x^\gamma: \gamma > 0)$ is a deterministic sequence with $\lim_{\gamma\to\infty} x^\gamma = x_o$. Then the sequence $(\tilde{X}^\gamma: \gamma > 0)$ satisfies the large deviations principle in $C_{[0,T]}(R^d)$ with the good rate function $\Gamma(\cdot, x_o)$, where for each $\phi \in C_{[0,T]}(R^d)$ and each $x_0 \in R^d$,

$$
\Gamma(\phi, x_0) = \begin{cases} \displaystyle\int_0^T \Lambda(\phi_t, \dot{\phi}_t)\,dt, & \text{if } \phi_0 = x_0 \text{ and } \phi \\ & \quad \text{is absolutely continuous}, \\ +\infty, & \text{otherwise}, \end{cases}
\tag{2.3}
$$

and Λ *satisfies*

$$\Lambda(\phi_t, \dot{\phi}_t) = I\{\phi_t \in A^+\}\Lambda^+(\phi_t, \dot{\phi}_t) + I\{\phi_t \in A^o\}\Lambda^o(\phi_t, \dot{\phi}_t) \\ + I\{\phi_t \in A^-\}\Lambda^-(\phi_t, \dot{\phi}_t). \tag{2.4}$$

REMARK 2.1. Theorem 2.1 gives many large deviations principles for each fixed x_o, since the sequence of intial conditions (x^γ: $\gamma > 0$) can be varied. In the terminology of Dinwoodie and Zabell (1992), the collection of probability measures of $\tilde{X}^\gamma$ parameterized by the initial state x^γ is *exponentially continuous*. Let P_x denote a probability measure under which $\tilde{X}^\gamma$ has initial state x. A simple consequence of Theorem 2.1 is that the following inequalities hold for each $\phi \in C_{[0,T]}(R^d)$:

$$\lim_{\delta\searrow 0} \limsup_{\gamma\to\infty} \gamma^{-1} \log \sup_{|x-\phi_0|<\delta} P_x\Big(\sup_{0\le t\le T} |\tilde{X}^\gamma_t - \phi_t| < \delta\Big) \le -\Gamma(\phi, \phi_0),$$

$$\lim_{\delta\searrow 0} \lim_{\rho\searrow 0} \liminf_{\gamma\to\infty} \gamma^{-1} \log \inf_{|x-\phi_0|<\rho} P_x\Big(\sup_{0\le t\le T} |\tilde{X}^\gamma_t - \phi_t| < \delta\Big) \ge -\Gamma(\phi, \phi_0).$$

REMARK 2.2 (Alternative representation of Λ^o). Let $n = (1, 0, \dots, 0) \in R^d$, and define

$$\tilde{\Lambda}^+(x, y) = \begin{cases} \Lambda^+(x, y), & \text{if } y \in \overline{A^-}, \\ +\infty, & \text{otherwise.} \end{cases}$$

Then

$$\Lambda^o(x, y) = \inf_{\substack{0\le\beta\le 1,\ y^+\in R^d,\ y^-\in R^d \\ \beta y^+ + (1-\beta)y^- = y}} \big\{\beta\tilde{\Lambda}^+(x, y^+) + (1-\beta)\Lambda^-(x, y^-)\big\}.$$

It is easy to check that $\tilde{\Lambda}^+(x, \cdot)$ is the Legendre–Fenchel transform of $\inf_{\alpha\ge 0} M^+(x, \cdot - \alpha n)$, therefore by Theorem 16.5 of Rockafellar (1970), $\Lambda^o(x, \cdot)$ is the Legendre–Fenchel transform of $\inf_{\alpha\ge 0} M^+(x, \cdot - \alpha n) \vee M^-(x, \cdot)$. In particular, for $y \in A^o$,

$$\Lambda^o(x, y) = \sup_{\zeta(2),\dots,\zeta(d)} \Big\{ y\zeta - \inf_{u\le v} M^-(x, (v, \zeta(2), \dots, \zeta(d))) \\ \vee M^+(x, (u, \zeta(2), \dots, \zeta(d)))\Big\}.$$

Note also that if $\nu^+(x, A^+) = 0$, (as in the case of a process in $\overline{A^-}$ with a flat boundary which cannot be crossed) then $\inf_{\alpha\ge 0} M^+(x, \zeta - \alpha n) = M^+(x, \zeta)$ and

$$\Lambda^o(x, y) = \sup_{\zeta\in R^d} \{y\zeta - M^-(x, \zeta) \vee M^+(x, \zeta)\},$$

as found in Shwartz and Weiss (1995).

The proof of Theorem 2.1 can be easily adapted to yield the following theorem for discrete time Markov chains.

THEOREM 2.2. *Let ν^+ and ν^- be two probability-measure fields on R^d each of which satisfies Conditions* 2.2 *and* 2.3, *and $\nu^+(x, A^-) > 0$ and $\nu^-(x, A^+) > 0$ for some (equivalently all) $x \in R^d$. For $\gamma > 0$, let $(X_k^\gamma\colon k \in Z_+)$ denote the Markov chain such that given X_k^γ, the scaled increment $\gamma(X_{k+1}^\gamma - X_k^\gamma)$ has distribution $\nu(X_k^\gamma)$ where*

$$\nu(x) = \begin{cases} \nu^+(x), & \text{if } x \in \overline{A^+}, \\ \nu^-(x), & \text{if } x \in A^-, \end{cases}$$

and let $\tilde{X}^\gamma$ denote the polygonal interpolation of the process $(X_{\lfloor \gamma t \rfloor}^\gamma\colon t \geq 0)$. Suppose $\tilde{X}_0^\gamma = x^\gamma$ where $(x^\gamma\colon \gamma > 0)$ is a deterministic sequence with $\lim_{\gamma\to\infty} x^\gamma = x_o$. Then the sequence $(\tilde{X}^\gamma\colon \gamma > 0)$ satisfies the large deviations principle in $C_{[0,T]}(R^d)$ with the good rate function $\Gamma(\cdot, x_o)$, where Λ^+, Λ^-, Λ^o are defined by (2.1) *and* (2.2) *with $M^\pm(x, \zeta) = \log \int_{R^d} \exp(z\zeta)\nu^\pm(x, dz)$.*

The proof of Theorem 2.1 is organized as follows: Section 3 contains some observations which are instrumental for the proof. Section 4 extends the work of BD to continuous time random walks, hence proves the theorem in the special case of constant transition mechanisms in each half-space. In view of this, Sections 5 and 6 establish, respectively, the large deviations lower and upper bounds in the general case. Goodness of the rate function is shown in Section 7.

3. Preliminaries. This section contains preliminary results regarding the proof of Theorem 2.1. Lemma 3.2 establishes that the process X^γ and its polygonal interpolation $\tilde{X}^\gamma$ are close in a certain sense, so that they have equivalent large deviation probabilities. The section concludes with Lemma 3.3 on the sensitivity of the rate function to variations of the rate measures.

LEMMA 3.1. *Given $x \in R^d$ and $\gamma > 0$, let $\gamma\Delta_x$ have distribution $\nu(x)/\nu(x, R^d)$. Then for each $\delta > 0$, $\limsup_{\gamma\to\infty} \gamma^{-1} \log \sup_x P(|\Delta_x| \geq \delta) = -\infty$.*

PROOF. If $|\Delta_x| \geq \delta$, then for some coordinate $1 \leq i \leq d$, $|\Delta_x(i)| \geq \delta/\sqrt{d}$. This, together with the union bound and Chernoff's inequality imply that

$$\sup_x P(|\Delta_x| \geq \delta) \leq 2d \exp(-\alpha\gamma\delta/\sqrt{d}) \sup_{x, 1\leq i\leq d} E[\exp(\alpha\gamma\Delta_x(i))] \quad \text{for each } \alpha \geq 0.$$

By Condition 2.2, $\sup_{x,1\leq i\leq d} E[\exp(\alpha\gamma\Delta_x(i))]$ is finite and it does not depend on γ, so that

$$\limsup_{\gamma\to\infty} \gamma^{-1} \log \sup_x P(|\Delta(x)| \geq \delta) \leq -\alpha\delta/\sqrt{d}.$$

The arbitrariness of $\alpha \geq 0$ yields the desired result. □

LEMMA 3.2 (Exponential equivalence). *For each $\delta > 0$,*

$$\limsup_{\gamma\to\infty} \gamma^{-1} \log \sup_x P_x\Big(\sup_{0\leq t\leq T} |X_t^\gamma - \tilde{X}_t^\gamma| > \delta \Big) = -\infty.$$

PROOF. Let N_T^γ denote the number of jumps of X^γ in the interval $[0, T]$. Note that if $\sup_{0\le t\le T}|X_t^\gamma - -\tilde{X}_t^\gamma| > \delta$ then at least one of the first $N_T^\gamma + 1$ jumps of X^γ has size larger then δ. Therefore for each $\gamma > 0$, $B > 0$ and $x \in R^d$,

$$\begin{aligned} P_x\Big(\sup_{0\le t\le T}|X_t^\gamma - \tilde{X}_t^\gamma| > \delta\Big) \\ &\le P_x(N_T^\gamma \ge \gamma B) + P_x\Big(\sup_{0\le t\le T}|X_t^\gamma - \tilde{X}_t^\gamma| > \delta,\ N_T^\gamma < \gamma B\Big) \\ &\le P_x(N_T^\gamma \ge \gamma B) + (\gamma B + 1)\sup_{x'} P(|\Delta_{x'}| \ge \delta), \end{aligned}$$

where $\gamma\Delta_x$ has distribution $\nu(x)/\nu(x, R^d)$. By Condition 2.1, uniformly over all initial states, N_T^γ is stochastically dominated by a Poisson random variable with mean $\gamma m T$. Therefore, given $K > 0$, B can be taken large enough so that

$$\begin{aligned} \limsup_{\gamma\to\infty}\gamma^{-1}\log\sup_x P_x\Big(\sup_{0\le t\le T}|X_t^\gamma - \tilde{X}_t^\gamma| > \delta\Big) \\ &\le \Big(\limsup_{\gamma\to\infty}\gamma^{-1}\log\sup_x P_x(N_T^\gamma \ge \gamma B)\Big) \\ &\qquad \bigvee\Big(\limsup_{\gamma\to\infty}\gamma^{-1}\log\sup_x P(|\Delta_x| \ge \delta)\Big) \qquad (3.1) \\ &= \limsup_{\gamma\to\infty}\gamma^{-1}\log\sup_x P_x(N_T^\gamma \ge \gamma B) \\ &\le -K, \end{aligned}$$

where (3.1) follows by Lemma 3.1. The arbitrariness of $K > 0$ proves the lemma. □

Given a Borel measure μ on R^d, define

$$\Lambda_\mu(y) = \sup_{\zeta\in R^d}\Big\{y\zeta - \int_{R^d}(e^{z\zeta} - 1)\mu(dz)\Big\}, \qquad y \in R^d.$$

LEMMA 3.3. *If ν_0 and ν_1 are two positive, finite Borel measures on R^d such that $(1+\varepsilon)^{-1} \le d\nu_0/d\nu_1 \le (1+\varepsilon)$ for some $\varepsilon > 0$, then for all $y \in R^d$,*

$$\Lambda_{\nu_0}(y) \ge (1+\varepsilon)^{-1}\Lambda_{\nu_1}(y) - \varepsilon e\nu_1(R^d).$$

PROOF. Define $\chi(\varepsilon) = \sup_{u\in R}\{(1+\varepsilon)^2 e^u - e^{(1+\varepsilon)u}\}$. Straightforward evaluation yields that $\chi(\varepsilon) = \varepsilon(1+\varepsilon)^{(1+\varepsilon)/\varepsilon}$. For each $\zeta \in R^d$,

$$\begin{aligned} \int_{R^d}(e^{z\zeta} - 1)\nu_0(dz) &= \int_{R^d} e^{z\zeta}\nu_0(dz) - \nu_0(R^d) \qquad (3.2) \\ &\le (1+\varepsilon)\int_{R^d} e^{z\zeta}\nu_1(dz) - (1+\varepsilon)^{-1}\nu_1(R^d) \end{aligned}$$

$$\leq (1+\varepsilon)^{-1}\int_{R^d} e^{z\zeta(1+\varepsilon)}\nu_1(dz) \qquad (3.3)$$
$$+ (\chi(\varepsilon)-1)(1+\varepsilon)^{-1}\nu_1(R^d)$$
$$\leq (1+\varepsilon)^{-1}\int_{R^d} (e^{z\zeta(1+\varepsilon)}-1)\nu_1(dz) + \varepsilon e\nu_1(R^d), \qquad (3.4)$$

where inequality (3.2) is a consequence of the hypothesis, (3.3) is implied by the definition of $\chi(\varepsilon)$, and (3.4) follows by the fact that $\chi(\varepsilon)/(1+\varepsilon) \leq \varepsilon e$. This in turn implies that for any $y \in R^d$,

$$\begin{aligned}
\Lambda_{\nu_0}(y) &= \sup_{\zeta\in R^d}\left\{ y\zeta - \int_{R^d}(e^{z\zeta}-1)\nu_0(dz)\right\} \\
&\geq \sup_{\zeta\in R^d}\left\{ y\zeta - (1+\varepsilon)^{-1}\int_{R^d}(e^{z\zeta(1+\varepsilon)}-1)\nu_1(dz)\right\} - \varepsilon e\nu_1(R^d) \\
&= (1+\varepsilon)^{-1}\sup_{\zeta\in R^d}\left\{ y\zeta(1+\varepsilon) - \int_{R^d}(e^{z\zeta(1+\varepsilon)}-1)\nu_1(dz)\right\} - \varepsilon e\nu_1(R^d) \\
&= (1+\varepsilon)^{-1}\Lambda_{\nu_1}(y) - \varepsilon e\nu_1(R^d).
\end{aligned}$$

This proves the lemma. □

4. The piecewise homogeneous case. This section establishes Theorem 2.1 in case the two rate-measure fields are constant. The result, stated as Lemma 4.1 below, can be proved by adapting the proof of the analogous result for discrete time piecewise homogeneous random walks, as presented in BD. We shall here outline the sufficient modifications of that proof, while at the same time pointing out how the argument in BD can be somewhat streamlined.

LEMMA 4.1. *If $\nu^+, \nu^-, \tilde{X}^\gamma$ satisfy the conditions of Theorem 2.1 with $\nu^+(x) \equiv \nu_o^+$, $\nu^-(x) \equiv \nu_o^-$ for two fixed measures ν_o^+ and ν_o^-, then the sequence $(\tilde{X}^\gamma\colon \gamma > 0)$ satisfies the large deviations principle with the good rate function $\Gamma(\cdot, x_o)$.*

Let $(s_t^\pm\colon t \geq 0)$ denote a compound Poisson process with rate measure $\nu_o^\pm$, so that the probability distribution $P^\pm$ of the random variable $s_1^\pm$ is a compound Poisson probability distribution with log moment generating function $G_P^\pm$ given by

$$G_P^\pm(\zeta) = \log\int_{R^d} e^{z\zeta}P^\pm(dz) = \int_{R^d}(e^{z\zeta}-1)\nu_o^\pm(dz) = M^\pm(\zeta).$$

(The first arguments of $M^\pm$ and $\Lambda^\pm$ are suppressed in this section, since the rate-measure fields are constant.) Thus, $\Lambda^\pm$ is the Legendre–Fenchel transform (denoted $H_P^\pm$ in BD) of $G_P^\pm$. The expression $\Gamma(\phi, x_0)$ of Theorem 2.1 is thus identical to the rate function $N(\phi)$ defined in BD, hence we simply refer to that paper for the properties of $\Lambda^\pm$, Λ^o and Γ. In particular, it is shown there that Γ is a good rate function.

A representation of $\tilde{X}^\gamma$. The key to the proof in BD is to combine two independent homogeneous random walks to produce a single, piecewise homogeneous random walk. The continuous time process X^γ can similarly be constructed by combining the processes s^+ and s^-, as shown below. The random variables $s_t^\pm/t$ obey the Cramér theorem as $t \to +\infty$, with the rate function $\Lambda^\pm$. Furthermore, the exponential tightness property and local large deviation properties hold exactly as for the discrete time case stated in Proposition 5.3 of BD.

We next define an "unscaled" process X so that the process X^γ has the same distribution as the process $((X_{\gamma t})/\gamma\colon t \geq 0)$. The process X is conveniently defined via a jump representation, using the following jump representations of $s^\pm$. Let $(J^\pm(k)\colon k \geq 1)$ be independent, identically distributed random variables with the probability distribution $\nu_o^\pm(\cdot)/\nu_o^\pm(R^d)$. Let $(U^\pm(k)\colon k \geq 1)$ be independent, exponentially distributed random variables with parameter $\nu_o^\pm(R^d)$. Also, for convenience, set $J^\pm(0) = U^\pm(0) = 0$. Then $s^\pm$ can be represented as

$$s_t^\pm = J^\pm(0) + \cdots + J^\pm(k) \quad \text{if } U^\pm(0) + \cdots + U^\pm(k) \leq t < U^\pm(0) + \cdots + U^\pm(k+1).$$

Of course it is assumed that s^+ is independent of s^-. Given an initial state X_0, let X denote the Markov process for which the corresponding variables $(U(k)\colon k \geq 0)$ and $(J(k)\colon k \geq 0)$ are defined recursively as follows. $U(0) = 0$, $n^\pm(0) = 0$, $J(0) \equiv X_0$ and

$$\text{if } X_0 + J(1) + \cdots + J(k) \in \overline{A^+} \text{ then}$$
$$\begin{cases} n^+(k+1) = n^+(k) + 1, & U(k+1) = U^+(n^+(k+1)), \\ n^-(k+1) = n^-(k), & J(k+1) = J^+(n^+(k+1)), \end{cases}$$

$$\text{else if } X_0 + J(1) + \cdots + J(k) \in A^- \text{ then}$$
$$\begin{cases} n^+(k+1) = n^+(k), & U(k+1) = U^-(n^-(k+1)), \\ n^-(k+1) = n^-(k) + 1, & J(k+1) = J^-(n^-(k+1)). \end{cases}$$

Then X can be represented as

$$X_t = X_0 + J(1) + \cdots + J(k) \quad \text{if } U(0) + \cdots + U(k) \leq t < U(0) + \cdots + U(k+1).$$

Note that the process $((X_{\gamma t})/\gamma\colon t \geq 0)$ with $X_0 = \gamma x^\gamma$ can be identified with the process X^γ as desired. Define $(\Theta_t\colon t \geq 0)$ as follows. If $U(0) + \cdots + U(k) \leq t < U(0) + \cdots + U(k+1)$, then

$$\Theta_t = \begin{cases} 1, & \text{if } X_0 + J(1) + \cdots + J(k) \in \overline{A^+}, \\ 0, & \text{else.} \end{cases}$$

Intuitively, X evolves according to s^+ on the intervals in which $\Theta_t = 1$. In particular, let $\tau(t) = \int_0^t \Theta_s \, ds$. Then for $t \geq 0$,

$$X_t = X_0 + s^+_{\tau(t)} + s^-_{t-\tau(t)}. \tag{4.1}$$

Identify $\tilde{X}^\gamma$ as the polygonal interpolation of the scaled process $((X_{\gamma t})/\gamma\colon t \geq 0)$, assuming $X_0 = \gamma x^\gamma$. Let S^+ and S^- denote the polygonal interpolations of $((s^+_{\gamma t})/\gamma\colon t \geq 0)$ and $((s^-_{\gamma t})/\gamma\colon t \geq 0)$ respectively (for brevity we do not explicitly indicate the dependence of $S^\pm$ on γ). It is useful to note that the relation (4.1) carries over to the scaled processes

$$\tilde{X}^\gamma_t = x^\gamma + S^+_{\tau(t)} + S^-_{t-\tau(t)}, \tag{4.2}$$

for all $t \geq 0$.

Given $\eta > 0$ and $T > 0$, define the events

$$K^\pm(\eta, T, \gamma) = \{|S^\pm_t - (s^\pm_{\gamma t})/\gamma| \leq \eta,\ 0 \leq t \leq T\}$$

and set $K(\eta, T, \gamma) = K^+(\eta, T, \gamma) \cap K^-(\eta, T, \gamma)$. Lemma 3.2 implies that the set $K(\eta, T, \gamma)^c$ is negligible for the purposes of proving large deviations principles, in the sense that

$$\lim_{\gamma\to\infty} \gamma^{-1} \log P[K(\eta, T, \gamma)^c] = -\infty.$$

Note that on the event $K(\eta, T, \gamma)$, $|\tilde{X}^\gamma_t - (X_{\gamma t})/\gamma| \leq \eta \quad$ for $0 \leq t \leq T$.

Due to the analytic considerations in BD, the proof of the large deviations principle in continuous time can be reduced to proving upper and lower large deviations bounds for the events of the form $\mathscr{E}(\sigma, \delta, \gamma, T) = \{|\tilde{X}^\gamma_t - \sigma_t| \leq \delta,\ 0 \leq t \leq T\}$, where $T > 0$ and $\delta > 0$ can be taken arbitrarily small. Here $\sigma_t = x_0 + tv$, where $v \in R^d$ and, with $x_1 = \sigma_T$, either $x_0, x_1 \in \overline{A^+}$ or $x_0, x_1 \in \overline{A^-}$. The key to proving these bounds is to bound the event $\mathscr{E}(\sigma, \delta, \gamma, T)$ from inside and outside by simple events involving the process (S^+, S^-), and to appeal to the large deviations principle for (S^+, S^-). This is essentially the same idea as in BD, translated for continuous time. Our proof is simplified somewhat in that (1) in the case of the upper bound, our proof makes better use of the large deviations principle for (S^+, S^-), (which is common to both discrete and continuous time), and (2) we exploit the representation (4.2). These simplifications make the translation between discrete and continuous time more transparent.

Lower bound. The three lemmas that follow identify events involving (S^+, S^-) which are subsets of the event $\mathscr{E}(\sigma, \delta, \gamma, T)$ whenever $x_0, x_1 \in \overline{A^+}$. The case $x_0, x_1 \in \overline{A^-}$ can be handled similarly. The large deviations lower bound for the process (S^+, S^-) can then be readily used to provide the required lower bound for $P[\mathscr{E}(\sigma, \delta, \gamma, T)]$.

LEMMA 4.2. *If $x_0, x_1 \in A^+$, then for δ small enough*,

$$\mathscr{E}(\sigma, \delta, \gamma, T) \supset \{|x^\gamma - x_0| \leq \delta/2\} \cap \{|S^+_t - tv| \leq \delta/2,\ 0 \leq t \leq T\}. \tag{4.3}$$

PROOF. It is enough to note that for δ small enough, $\tilde{X}^\gamma_t = x^\gamma + S^+_t$ for $0 \leq t \leq T$ if the event on the right-hand side of (4.3) is true. □

Corollary 3.2 of BD states that there is a vector $b^- \in A^+$ such that $\Lambda^-(b^-) < +\infty$.

LEMMA 4.3. *If* $(x_0 \in A^+,\ x_1 \in A^o)$ *or if* $(x_0 \in A^o,\ x_1 \in A^+)$ *then there exist* $\eta = \eta(\delta) \to 0$ *and* $\kappa = \kappa(\delta) \to 0$ *as* $\delta \to 0$ *so that*

$$\begin{aligned}\mathscr{E}(\sigma,\delta,\gamma,T) \supset \{&|S_t^+ - tv| \le \eta,\ 0 \le t \le T\}\\ &\cap \{|S_t^- - tb^-| \le \eta,\ 0 \le t \le \kappa\}\\ &\cap \{|x^\gamma - x_0| \le \eta\} \cap K(\eta, T, \gamma).\end{aligned} \tag{4.4}$$

PROOF. Assume that $(x_0 \in A^o,\ x_1 \in A^+)$. Take $\kappa = 5\eta/b^-(1)$, where $\eta = \eta(\delta)$ is yet to be specified, and suppose that the event on the right-hand side of (4.4) is true. We first prove the following claim: $T - \tau(T) < \kappa$. If this claim is false, let u denote the minimum positive value such that $u - \tau(u) = \kappa$. Then $\kappa \le u \le T$ and

$$\begin{aligned}\tilde{X}_u^\gamma(1) &= x^\gamma(1) + S^+_{\tau(u)}(1) + S^-_{u-\tau(u)}(1)\\ &\ge -\eta + (u-\kappa)v(1) - \eta + \kappa b^-(1) - \eta\\ &\ge \kappa b^-(1) - 3\eta = 2\eta.\end{aligned}$$

On the event $K(\eta, T, \gamma)$, $\Theta_t = 1$ whenever $\tilde{X}_t^\gamma(1) \ge \eta$ so that u cannot be a point of increase of $t - \tau(t)$. The claim is thus true by proof by contradiction.

Thus, for $0 \le t \le T$,

$$\begin{aligned}|\tilde{X}_t^\gamma - \sigma_t| &\le |\tilde{X}_t^\gamma - (x_0 + S_t^+)| + |x_0 + S_t^+ - \sigma_t|\\ &\le |x_0 - x^\gamma| + \Big(\sup_{t-\kappa\le r\le t} |S_r^+ - S_t^+|\Big) + \Big(\sup_{0\le r\le\kappa} |S_r^-|\Big) + |tv - S_t^+|\\ &\le \eta + (|v|\kappa + 2\eta) + (|b^-|\kappa + \eta) + \eta = C\eta,\end{aligned}$$

where the constant C depends only on v and b^-. Taking $\eta = \delta/C$, the event $\mathscr{E}(\sigma, \delta, \gamma, t)$ is true, and the lemma is proved in the case $(x_0 \in A^+,\ x_1 \in A^o)$. The proof in the case $(x_0 \in A^o,\ x_1 \in A^+)$ is similar and is omitted. □

LEMMA 4.4. *If* $x_0, x_1 \in A^o$, *and if* $0 < \beta < 1$, $v^+ \in A^-$ *and* $v^- \in A^+$ *are such that* $v = \beta v^+ + (1-\beta)v^-$, *then there exist* $\eta = \eta(\delta) \to 0$ *and* $\kappa = \kappa(\delta) \to 0$ *as* $\delta \to 0$ *so that*

$$\begin{aligned}\mathscr{E}(\sigma,\delta,\gamma,T) \supset \{&|S_t^+ - tv^+| \le \eta,\ 0 \le t \le \beta T + \kappa\}\\ &\cap \{|S_t^- - tv^-| \le \eta, 0 \le t \le (1-\beta)T + \kappa\}\\ &\cap \{|x^\gamma - x_0| \le \eta\} \cap K(\eta, T, \gamma).\end{aligned} \tag{4.5}$$

PROOF. Take $\kappa = 5\eta/(v^-(1) - v^+(1))$, where $\eta = \eta(\delta)$ is yet to be specified, and suppose that the event on the right-hand side of (4.5) is true. We first prove the following claim: $|\tau(t) - \beta t| < \kappa$, or equivalently, $|t - \tau(t) - (1-\beta)t| < \kappa$, for $0 \le t \le T$. If this claim is false, let u denote the minimum value of t such that the inequalities are violated. Then either $\tau(u) = \beta u + \kappa$ or $u - \tau(u) = (1-\beta)u + \kappa$. By symmetry we assume without loss of generality that $u - \tau(u) = (1-\beta)u + \kappa$, and hence also $\tau(u) = \beta u - \kappa$. Thus,

$$\begin{aligned}\tilde{X}^{\gamma}_{u}(1) &= x^{\gamma}(1) + S^{+}_{\tau(u)}(1) + S^{-}_{u-\tau(u)}(1)\\ &\geq -\eta + (\beta u - \kappa)v^{+}(1) - \eta + ((1-\beta)u + \kappa)v^{-}(1) - \eta\\ &= \kappa(v^{-}(1) - v^{+}(1)) - 3\eta = 2\eta.\end{aligned}$$

On the event $K(\eta, t, \gamma)$, $\Theta_t = 1$ whenever $\tilde{X}^{\gamma}_{t}(1) \geq \eta$ so that u cannot be a point of increase of $t - \tau(t)$. The claim is thus true by proof by contradiction.

Thus, for $0 \leq t \leq T$,

$$\begin{aligned}|\tilde{X}^{\gamma}_{t} - \sigma_t| &\leq |\tilde{X}^{\gamma}_{t} - (x_0 + S^{+}_{\beta t} + S^{-}_{(1-\beta)t})| + |S^{+}_{\beta t} - \beta t v^{+}|\\ &\quad + |S^{-}_{(1-\beta)t} - (1-\beta)tv^{-}|\\ &\leq |x^{\gamma} - x_0| + \Big(\sup_{|r-\beta t|\leq\kappa} |S^{+}_{r} - S^{+}_{\beta t}|\Big)\\ &\quad + \Big(\sup_{|r-(1-\beta)t|\leq\kappa} |S^{-}_{r} - S^{-}_{(1-\beta)t}|\Big) + \eta + \eta\\ &\leq \eta + (2|v^{+}|\kappa + 2\eta) + (2|v^{-}|\kappa + 2\eta) + 2\eta = C\eta,\end{aligned}$$

where the constant C depends only on v^{+} and v^{-}. Taking $\eta = \delta/C$, the event $\mathscr{E}(\sigma, \delta, \gamma, t)$ is true, and the lemma is proved. $\square$

Proposition 3.4 of BD shows that the conditions $y^{\pm} \in \overline{A^{\pm}}$ and $0 \leq \beta \leq 1$ in (2.2), can be replaced by the conditions $y^{\pm} \in A^{\pm}$ and $0 < \beta < 1$, without changing the value of Λ^{o}. Thus Lemma 4.4, with its condition that $v^{\pm} \in A^{\pm}$ (rather than $v^{\pm} \in \overline{A^{\pm}}$) and $0 < \beta < 1$ suffices for the derivation of the required lower large deviations bound for $\mathscr{E}(\sigma, \delta, \gamma, T)$.

Upper bound. The two lemmas that follow identify events involving (S^{+}, S^{-}) which contain the event $\mathscr{E}(\sigma, \delta, \gamma, T)$, whenever $x_0, x_1 \in \overline{A^{+}}$. The case $x_0, x_1 \in \overline{A^{-}}$ can be handled similarly. The large deviations upper bound for the process (S^{+}, S^{-}) can then be readily used to provide the required upper bound for $P[\mathscr{E}(\sigma, \delta, \gamma, T)]$. The first lemma is easily verified, and is stated without proof.

LEMMA 4.5. *If $x_0, x_1 \in \overline{A^{+}}$, and $\{x_0, x_1\} \not\subset A^{o}$, then there exist $\eta = \eta(\delta) \to 0$ and $\kappa = \kappa(\delta) \to 0$ as $\delta \to 0$ so that*

$$\mathscr{E}(\sigma, \delta, \gamma, T) \subset \{|S^{+}_{T-\kappa} - (T-\kappa)v| \leq \eta\} \cup K(\eta, T, \gamma)^{c}.$$

LEMMA 4.6. *If $x_0, x_1 \in A^{o}$, then for $\delta,\ \kappa,\ \varepsilon > 0$, $\mathscr{E}(\sigma, \delta, \gamma, T) \subset \{(S^{+}, S^{-}) \in F^{2\delta}\}$, where $F^{2\delta}$ and F are the subsets of $C_{[0,T]}(R^{d} \times R^{d})$ defined as follows*:

$$\begin{aligned}F^{2\delta} = \Big\{(\tilde{\phi}^{+}, \tilde{\phi}^{-})\colon &\sup_{0\leq t\leq T} |\tilde{\phi}^{+}_{t} - \phi^{+}_{t}| \leq 2\delta \text{ and}\\ &\sup_{0\leq t\leq T} |\tilde{\phi}^{-}_{t} - \phi^{-}_{t}| \leq 2\delta \text{ for some } (\phi^{+}, \phi^{-}) \in F\Big\}\end{aligned}$$

and F denotes the closed set $F_1 \cup F_2 \cup F_3 \cup F_4$ where

$$F_1 = \Big\{(\phi^+, \phi^-)\colon \sup_{0 \le t \le \kappa} (|\phi_t^+| + |\phi_t^-|) \ge \varepsilon\Big\},$$

$$F_2 = \big\{(\phi^+, \phi^-)\colon \exists \tau \in [\kappa, T-\kappa]\colon \phi_\tau^+ \in \overline{A^-}, \phi_{T-\tau}^- \in \overline{A^+}, \phi_\tau^+ + \phi_{T-\tau}^- = vT\big\},$$

$$F_3 = \big\{(\phi^+, \phi^-) : |\phi_{T-\kappa}^+ - (T-\kappa)v| \le \varepsilon + |v|\kappa\big\},$$

$$F_4 = \big\{(\phi^+, \phi^-) : |\phi_{T-\kappa}^- - (T-\kappa)v| \le \varepsilon + |v|\kappa\big\}.$$

PROOF. Suppose the event $\mathscr{E}(\sigma, \delta, \gamma, T)$ is true. Since

$$\mathscr{E}(\sigma, \delta, \gamma, T) = \big\{|x^\gamma + S_{\tau(t)}^+ + S_{t-\tau(t)}^- - \sigma_t| \le \delta, 0 \le t \le T\big\}$$

it follows (take $t = 0$) that $|x^\gamma - x_0| \le \delta$, so that

$$|S_{\tau(t)}^+ + S_{t-\tau(t)}^- - vt| \le 2\delta, \qquad 0 \le t \le T. \tag{4.6}$$

To complete the proof of the lemma we consider three cases.

Case 1. Suppose $\kappa \le \tau(T) \le T - \kappa$. Let $\triangle = S_{\tau(T)}^+ + S_{T-\tau(T)}^- - vT$, and note that $|\triangle| \le 2\delta$. Note by the construction of the process Y, if $\Theta_T = 0$ then $S_{\tau(T)}^+ \in \overline{A^-}$, whereas if $\Theta_T = 1$ then $S_{T-\tau(T)}^- \in \overline{A^+}$. Define (ϕ^+, ϕ^-) by setting

$$(\phi_t^+, \phi_t^-) = \begin{cases} \left(S_t^+, S_t^- - \triangle\left(\dfrac{t}{T-\tau(T)} \wedge 1\right)\right), & \text{if } \Theta_T = 0, \\ \left(S_t^+ - \triangle\left(\dfrac{t}{\tau(T)} \wedge 1\right), S_t^-\right), & \text{if } \Theta_T = 1, \end{cases}$$

for $0 \le t \le T$. Then $(\phi^+, \phi^-) \in F$ and $\sup_{0 \le t \le T} |S_t^\pm - \phi_t^\pm| \le 2\delta$ so that $(S^+, S^-) \in F^{2\delta}$.

Case 2. Suppose $\tau(T) > T - \kappa$. Let $t_0 = \min\{t \ge 0\colon \tau(t) = T - \kappa\}$ and let $t_1 = t_0 - (T - \kappa)$. Then $T - \kappa \le t_0 \le T$ and $0 \le t_1 \le \kappa$. Also, $\tau(t_0) = T - \kappa$ and $t_0 - \tau(t_0) = t_1$, so by (4.6),

$$\big|S_{T-\kappa}^+ + S_{t_1}^- - (t_1 + T - \kappa)v\big| \le 2\delta. \tag{4.7}$$

We assume in addition that

$$\sup_{0 \le t \le \kappa} |S_t^-| \le \varepsilon, \tag{4.8}$$

for otherwise $(S^+, S^-) \in F \subset F^{2\delta}$. Combining (4.7), (4.8) and the fact $0 \le t_1 \le \kappa$ yields that

$$\big|S_{T-\kappa}^+ - (T-\kappa)v\big| \le 2\delta + \varepsilon + |v|\kappa.$$

Therefore, $(S^+, S^-) \in F^{2\delta}$.

Case 3. Suppose $\tau(T) \le \kappa$. This case is the same as Case 2 with the roles of S^+ and S^- reversed. Lemma 4.6 is thus proved. □

Lemma 4.5 immediately implies that if $x_0, x_1 \in \overline{A^+}$, and $\{x_0, x_1\} \not\subset A^o$, then

$$\lim_{\delta\to 0} \limsup_{\gamma\to\infty} \gamma^{-1} \log P[\mathscr{E}(\sigma, \delta, \gamma, T)] \le -T\Lambda^+(v).$$

Similarly, Lemma 4.6 yields the appropriate large deviations upper bound if $x_0, x_1 \in A^o$.

LEMMA 4.7. *If $x_0, x_1 \in A^o$, then*

$$\lim_{\delta\to 0} \limsup_{\gamma\to\infty} \gamma^{-1} \log P[\mathscr{E}(\sigma, \delta, \gamma, T)] \le -T\Lambda^o(v).$$

PROOF. Let $\Gamma^\pm$ be defined as Γ in (2.3), but with $\Lambda \equiv \Lambda^\pm$. The process (S^+, S^-) satisfies a large deviations principle with the good rate function $\Gamma^+(\cdot, 0) + \Gamma^-(\cdot, 0)$, so by Lemma 4.6 for each $\kappa,\ \varepsilon > 0$,

$$\begin{aligned}\lim_{\delta\to 0} \limsup_{\gamma\to\infty} \gamma^{-1} \log P[\mathscr{E}(\sigma, \delta, \gamma, T)] &\le -\lim_{\delta\to 0} \inf_{(\phi^+,\, \phi^-)\in F^{2\delta}} \{\Gamma^+(\phi^+, 0) + \Gamma^-(\phi^-, 0)\}\\ &= -\inf_{(\phi^+,\, \phi^-)\in F} \big\{\Gamma^+(\phi^+, 0) + \Gamma^-(\phi^-, 0)\big\}.\end{aligned}$$

To complete the proof of the lemma, it suffices to show that for each $\rho > 0$, there exist $\kappa, \varepsilon > 0$ such that for each $j \in \{1, 2, 3, 4\}$

$$\inf_{(\phi^+,\, \phi^-)\in F_j} \{\Gamma^+(\phi^+, 0) + \Gamma^-(\phi^-, 0)\} \ge T\Lambda^o(v) - \rho. \tag{4.9}$$

Note that inequality (4.9) holds for $j = 2$ for all $\rho, \varepsilon, \kappa > 0$. Choose $L > T\Lambda^o(v)$. The fact that $\Lambda^\pm(y)/|y| \to \infty$ as $|y| \to \infty$ implies the existence of $\kappa(\varepsilon) \to 0$ as such that for each ε,

$$\inf_{(\phi^+,\, \phi^-)\in F_1} \big\{\Gamma^+(\phi^+, 0) + \Gamma^-(\phi^-, 0)\big\} > L,$$

so that (4.9) holds for $j = 1$. Since

$$\begin{aligned}\inf_{(\phi^+,\, \phi^-)\in F_3} \big\{\Gamma^+(\phi^+, 0) + \Gamma^-(\phi^-, 0)\big\} &= \inf_{|y^+ - v| \le \varepsilon + |v|\kappa(\varepsilon)} (T - \kappa(\varepsilon))\Lambda^+(y^+)\\ &\to T\Lambda^+(v) \quad \text{as } \varepsilon \to 0\end{aligned}$$

and $\Lambda^+(v) \ge \Lambda^o(v)$, inequality (4.9) holds for $j = 3$, and similarly for $j = 4$, for sufficiently small ε. □

5. The lower bound. This section establishes the large deviations lower bound for Theorem 2.1, roughly as follows: given $\phi \in C_{[0,\, T]}(R^d)$, the process X^γ is approximated by a "patchwork" Markov process with a time varying transition mechanism that for each t is constant on each half-space. The time variation is determined by ϕ and a partition of $[0, T]$. This approach finds its roots in the work of Azencott and Ruget (1977). Lemma 4.1 is used to prove a local lower bound for the patchwork process. Then by comparing the quantities on each side of this inequality to the corresponding quantities for X^γ, a local lower bound is obtained for X^γ. Following standard techniques, this local lower bound is shown to imply the lower bound for Theorem 2.1.

For $T > 0$, a *partition* of the interval $[0, T]$ is a finite sequence $\theta = (\theta_0, \dots, \theta_{J(\theta)})$ such that $0 = \theta_0 < \theta_1 < \cdots < \theta_{J(\theta)} = T$. Given $\phi \in C_{[0,T]}(R^d)$ and a partition θ of $[0, T]$, let $X^{\gamma,\phi,\theta}$ denote a Markov process with a time-varying transition mechanism: For each $i \in \{0, \dots, J(\theta) - 1\}$, $X^{\gamma,\phi,\theta}$ in the time interval $(\theta_i, \theta_{i+1}]$ is generated by the pair (γ, ν_i), where for each $x \in R^d$ the rate measure $\nu_i(x)$ satisfies

$$\nu_i(x) = \begin{cases} \nu^+(\phi_{\theta_i}), & \text{if } x \in \overline{A^+}, \\ \nu^-(\phi_{\theta_i}), & \text{if } x \in A^-. \end{cases}$$

Also let $\Lambda^{\phi_{\theta_i}}$ denote the function Λ defined by (2.4) when $\nu^+(x) \equiv \nu^+(\phi_{\theta_i})$ and $\nu^-(x) \equiv \nu^-(\phi_{\theta_i})$.

LEMMA 5.1 (Intermediate lower bound). *For each $T > 0$, partition $\theta = (\theta_0, \dots, \theta_{J(\theta)})$ of $[0, T]$, and absolutely continuous $\phi \in C_{[0,T]}(R^d)$,*

$$\begin{aligned} &\lim_{\delta\searrow 0} \lim_{\rho\searrow 0} \liminf_{\gamma\to\infty} \gamma^{-1} \log \inf_{|x-\phi_0|<\rho} P_x\Big(\sup_{0\le t\le T} |X_t^{\gamma,\phi,\theta} - \phi_t| < \delta \Big) \\ &\qquad \ge - \sum_{i=0}^{J(\theta)-1} \int_{\theta_i}^{\theta_{i+1}} \Lambda^{\phi_{\theta_i}}(\phi_t, \dot\phi_t)\, dt. \end{aligned}$$

PROOF. We prove the lemma by induction on $J(\theta)$. Lemma 4.1, along with Remark 2.1 and Lemma 3.2, implies that the statement of the lemma holds whenever $J(\theta) = 1$. As the induction hypothesis, let $k \ge 1$ and suppose that the lemma holds for any $T > 0$ and partition θ of $[0, T]$ such that $J(\theta) = k$. Then $\forall \varepsilon > 0\ \exists\ \delta_k(\varepsilon) > 0$ such that $\forall\ \delta \in (0, \delta_k(\varepsilon))\ \exists\ \rho_k(\delta, \varepsilon)$ such that $\forall\ \rho \in (0, \rho_k(\delta, \varepsilon))\ \exists\ \gamma_k(\rho, \delta, \varepsilon)$ such that for $\gamma > \gamma_k(\rho, \delta, \varepsilon)$,

$$\begin{aligned} &\gamma^{-1} \log \inf_{|x-\phi_0|<\rho} P_x\Big(\sup_{0\le t\le \theta_k} |X_t^{\gamma,\phi,\theta} - \phi_t| < \delta \Big) \\ &\qquad \ge - \sum_{i=0}^{k-1} \int_{\theta_i}^{\theta_{i+1}} \Lambda^{\phi_{\theta_i}}(\phi_t, \dot\phi_t)\, dt - \varepsilon. \end{aligned} \tag{5.1}$$

By the time-homogeneous Markov property of the pair $(X^{\gamma,\phi,\theta}, \phi)$, $\forall\ \varepsilon > 0\ \exists\ \delta(\varepsilon) > 0$ such that $\forall\ \delta \in (0, \delta(\varepsilon))\ \exists\ \rho(\delta, \varepsilon)$ such that $\forall\ \rho \in (0, \rho(\delta, \varepsilon))\ \exists\ \gamma(\rho, \delta, \varepsilon)$ such that for $\gamma > \gamma(\rho, \delta, \varepsilon)$,

$$\begin{aligned} &\gamma^{-1} \log \inf_{|y-\phi_{\theta_k}|<\rho} P_x\Big(\sup_{\theta_k\le t\le \theta_{k+1}} |X_t^{\gamma,\phi,\theta} - \phi_t| < \delta \mid X_{\theta_k}^{\gamma,\phi,\theta} = y \Big) \\ &\qquad \ge - \int_{\theta_k}^{\theta_{k+1}} \Lambda^{\phi_{\theta_i}}(\phi_t, \dot\phi_t)\, dt - \varepsilon. \end{aligned} \tag{5.2}$$

To show that the claim holds for $J(\theta) = k+1$, fix $\varepsilon > 0$. For all $\delta \in (0, \delta(\varepsilon/2))$, $\alpha = (\delta \wedge \delta_k(\varepsilon/2) \wedge \rho(\delta, \varepsilon/2))/2$, $\rho \in (0, \rho_k(\alpha, \varepsilon/2))$ and $\gamma > \gamma(\alpha, \delta, \varepsilon/2) \vee \gamma_k(\rho, \alpha, \varepsilon/2)$,

$$
\begin{aligned}
\gamma^{-1} \log & \inf_{|x-\phi_0|<\rho} P_x\Big(\sup_{0\le t\le \theta_{k+1}} |X_t^{\gamma,\phi,\theta} - \phi_t| < \delta \Big) \\
&\ge \gamma^{-1} \log \inf_{|x-\phi_0|<\rho} P_x\Big(\sup_{0\le t\le \theta_k} |X_t^{\gamma,\phi,\theta} - \phi_t| < \alpha \Big) \\
&\quad + \gamma^{-1} \log \inf_{|y-\phi_{\theta_k}|<\alpha} P_x\Big(\sup_{\theta_k\le t\le \theta_{k+1}} |X_t^{\gamma,\phi,\theta} - \phi_t| < \delta \mid X_{\theta_k}^{\gamma,\phi,\theta} = y \Big) \\
&\ge -\sum_{i=0}^{k} \int_{\theta_i}^{\theta_{i+1}} \Lambda^{\phi_{\theta_i}}(\phi_t, \dot\phi_t)\, dt - \varepsilon,
\end{aligned}
$$

where the first step follows by the Markov property of $X^{\gamma,\phi,\theta}$ and the fact that $\alpha \le \delta$, and the second step follows by the statements (5.1) and (5.2) together with the choice of α. This completes the induction step, and establishes the lemma. □

Given $T > 0$, a partition θ of the interval $[0, T]$, $\phi \in C_{[0,T]}(R^d)$, and $x \in R^d$, let $P_x^{\gamma, T}$ and $P_x^{\gamma,\phi,\theta,T}$ denote respectively the probability distributions of $(X_t^\gamma\colon 0 \le t \le T)$ and $(X_t^{\gamma,\phi,\theta}\colon 0 \le t \le T)$, with $X_0^\gamma \equiv X_0^{\gamma,\phi,\theta} \equiv x$. Note that both measures are concentrated on the space of piecewise constant functions which take values in R^d, equal to x at time 0, are right continuous, and have finite number of jumps in $[0, T]$. There is a version D of the Radon–Nikodym derivative $dP_x^{\gamma,T}/dP_x^{\gamma,\phi,\theta,T}$, which satisfies for any such function ω,

$$
(5.3)\quad D(\omega) = \exp\Big(-\gamma \int_0^T (\nu(\omega_t, R^d) - \nu(\omega_{\ell(t)}, R^d))\, dt \Big) \prod_{k=1}^{N_T(\omega)} \frac{d\nu(\omega_{\tau_k^-})}{d\nu(\omega_{\ell(\tau_k)})}(\Delta\omega_k),
$$

where $N_T(\omega)$ denotes the number of jumps of ω in $[0, T]$, τ_k and $\Delta\omega_k$ denote, respectively, the time and size of the kth jump of ω, and $\ell(t) = \max\{\theta_i\colon \theta_i < t\}$.

LEMMA 5.2 (Local lower bound). *For each $\phi \in C_{[0,T]}(R^d)$ and $x_0 \in R^d$,*

$$
\lim_{\delta\searrow 0} \lim_{\rho\searrow 0} \liminf_{\gamma\to\infty} \gamma^{-1} \log \inf_{|x-x_0|<\rho} P_x\Big(\sup_{0\le t\le T} |X_t^\gamma - \phi_t| < \delta \Big) \ge -\Gamma(\phi, x_0).
$$

PROOF. We may take $\Gamma(\phi, x_0) < \infty$ so that ϕ is absolutely continuous and $\phi_0 = x_0$. Fix $\varepsilon > 0$. Since both ν^+ and ν^- satisfy Condition 2.3, there exists a $\delta > 0$ such that

$$
(1+\varepsilon)^{-1} \le \frac{d\nu^+(x)}{d\nu^+(x')} \le (1+\varepsilon) \quad\text{and}\quad (1+\varepsilon)^{-1} \le \frac{d\nu^-(x)}{d\nu^-(x')} \le (1+\varepsilon),
$$

whenever $|x-x'| < 2\delta$. Appeal to the uniform continuity of ϕ on $[0, T]$ to choose a partition $\theta = (\theta_0, \dots, \theta_{J(\theta)})$ of $[0, T]$ such that $\sup_{\theta_i \le t \le \theta_{i+1}} |\phi_t - \phi_{\theta_i}| < \delta$ for each $i \in \{0, \dots, J(\theta) - 1\}$. Then Lemma 3.3 applied to the definition of $\Lambda^{\phi_{\theta_i}}$ implies that

$$\begin{aligned} &\sum_{i=0}^{J(\theta)-1} \int_{\theta_i}^{\theta_{i+1}} \Lambda^{\phi_{\theta_i}}(\phi_t, \dot{\phi}_t)\, dt \\ &\quad \le \sum_{i=0}^{J(\theta)-1} \int_{\theta_i}^{\theta_{i+1}} ((1+\varepsilon)\Lambda^{\phi_t}(\phi_t, \dot{\phi}_t) + (1+\varepsilon)\varepsilon e m)\, dt \\ &\quad = (1+\varepsilon)\Gamma(\phi, x_0) + (1+\varepsilon)\varepsilon e m T. \end{aligned} \tag{5.4}$$

Let $N_T^{\gamma, \phi, \theta}$ and N_T^{γ} denote, respectively, the number of jumps of $X^{\gamma,\phi,\theta}$ and X^{γ} in the interval $[0, T]$. Appeal to Condition 2.1 to choose a B large enough so that for all $x \in R^d$,

$$\limsup_{\gamma \to \infty} \gamma^{-1} \log P_x(N_T^{\gamma, \phi, \theta} \ge \gamma B) \le -(1 + \Gamma(\phi, x_0)).$$

The choice of θ and equation (5.3) imply that for each $\gamma > 0$ and $x \in R^d$,

$$\begin{aligned} &P_x\Big(\sup_{0\le t\le T} |X_t^{\gamma, \phi, \theta} - \phi_t| < \delta\Big) \\ &\quad \le P_x\Big(\sup_{0\le t\le T} |X_t^{\gamma, \phi, \theta} - \phi_t| < \delta,\, N_T^{\gamma, \phi, \theta} < \gamma B\Big) + P_x(N_T^{\gamma, \phi, \theta} \ge \gamma B) \\ &\quad \le e^{\gamma\varepsilon(mT+B)} P_x\Big(\sup_{0\le t\le T} |X_t^{\gamma} - \phi_t| < \delta,\ N_T^{\gamma} < \gamma B\Big) + P_x(N_T^{\gamma, \phi, \theta} \ge \gamma B) \\ &\quad \le e^{\gamma\varepsilon(mT+B)} P_x\Big(\sup_{0\le t\le T} |X_t^{\gamma} - \phi_t| < \delta\Big) + P_x(N_T^{\gamma, \phi, \theta} \ge \gamma B), \end{aligned} \tag{5.5}$$

where the second step uses the fact that $\log(1+\varepsilon) \le \varepsilon$. Inequality (5.5), together with Lemma 5.1 and inequality (5.4) applied to the left-hand side of (5.5), and the choice of B imply that

$$\begin{aligned} &-(1+\varepsilon)\Gamma(\phi, x_0) - (1+\varepsilon)\varepsilon e m T \\ &\quad \le \Big(\varepsilon(mT+B) + \lim_{\delta \searrow 0} \lim_{\rho \searrow 0} \liminf_{\gamma\to\infty} \gamma^{-1} \log \inf_{|x-\phi_0|<\rho} P_x\Big(\sup_{0\le t\le T} |X_t^{\gamma} - \phi_t| < \delta\Big)\Big) \\ &\qquad \bigvee(-(1 + \Gamma(\phi, x_0))). \end{aligned}$$

The lemma follows by the arbitrariness of $\varepsilon > 0$. □

Given $\phi \in C_{[0,T]}(R^d)$ and $\delta > 0$, let $B(\phi, \delta)$ denote the open ball of radius δ around ϕ.

LEMMA 5.3 (Lower bound). *For any Borel measurable* $S \subset C_{[0,T]}(R^d)$, $x_0 \in R^d$, *and sequence* $(x^\gamma: \gamma > 0)$ *such that* $\lim_{\gamma\to\infty} x^\gamma = x_0$,

$$\liminf_{\gamma\to\infty} \gamma^{-1} \log P_{x^\gamma}\big((\tilde{X}_t^\gamma: 0 \le t \le T) \in S\big) \ge -\inf_{\phi\in S^o} \Gamma(\phi, x_0).$$

PROOF. Fix $\phi \in S^o$, and let $\delta' > 0$ be such that $B(\phi, \delta)$ is contained in S for all $\delta < \delta'$. Lemma 3.2 and Lemma 5.2 imply that

$$\begin{aligned}
&\liminf_{\gamma\to\infty} \gamma^{-1} \log P_{x^\gamma}\big((\tilde{X}_t^\gamma: 0 \le t \le T) \in S\big) \\
&\quad\ge \lim_{\delta\searrow 0} \lim_{\rho\searrow 0} \liminf_{\gamma\to\infty} \gamma^{-1} \log \inf_{|x-x_0|<\rho} P_x\Big(\sup_{0\le t\le T} |\tilde{X}_t^\gamma - \phi_t| < \delta\Big) \\
&\quad\ge \lim_{\delta\searrow 0} \lim_{\rho\searrow 0} \liminf_{\gamma\to\infty} \gamma^{-1} \log \inf_{|x-x_0|<\rho} P_x\Big(\sup_{0\le t\le T} |X_t^\gamma - \phi_t| < \delta\Big) \\
&\quad\ge -\Gamma(\phi, x_0).
\end{aligned}$$

Since $\phi \in S^o$ is arbitrary, the lemma follows. □

6. The upper bound. This section establishes the large deviations upper bound for Theorem 2.1 by adapting the methods of Section 5.

LEMMA 6.1 (Intermediate upper bound). *For each* $T > 0$, *partition* $\theta = (\theta_0, \dots, \theta_{J(\theta)})$ *of* $[0, T]$, *and absolutely continuous* $\phi \in C_{[0,T]}(R^d)$,

$$\begin{aligned}
&\lim_{\delta\searrow 0} \limsup_{\gamma\to\infty} \gamma^{-1} \log \sup_{|x-\phi_0|<\delta} P_x\Big(\sup_{0\le t\le T} |X_t^{\gamma,\phi,\theta} - \phi_t| < \delta\Big) \\
&\qquad \le -\sum_{i=0}^{J(\theta)-1} \int_{\theta_i}^{\theta_{i+1}} \Lambda^{\phi_{\theta_i}}(\phi_t, \dot{\phi}_t)\, dt.
\end{aligned} \tag{6.1}$$

Furthermore if ϕ *is not absolutely continuous, then the left-hand side of* (6.1) *equals* $-\infty$.

PROOF. By induction on $J(\theta)$, Lemma 4.1, along with Remark 2.1 and Lemma 3.2, implies that the statement of the lemma holds whenever $J(\theta) = 1$. As the induction hypothesis, let $k \ge 1$ and suppose that the lemma holds for any $T > 0$ and partition θ of $[0, T]$ such that $J(\theta) = k$. To show that the claim holds for $J(\theta) = k + 1$, note that by the Markov property of $X^{\gamma,\phi,\theta}$,

$$\begin{aligned}
&\gamma^{-1} \log \sup_{|x-\phi_0|<\delta} P_x\Big(\sup_{0\le t\le \theta_{k+1}} |X_t^{\gamma,\phi,\theta} - \phi_t| < \delta\Big) \\
&\quad\le \gamma^{-1} \log \sup_{|x-\phi_0|<\delta} P_x\Big(\sup_{0\le t\le \theta_k} |X_t^{\gamma,\phi,\theta} - \phi_t| < \delta\Big) \\
&\qquad + \gamma^{-1} \log \sup_{|y-\phi_{\theta_k}|<\delta} P_x\Big(\sup_{\theta_k\le t\le \theta_{k+1}} |X_t^{\gamma,\phi,\theta} - \phi_t| < \delta \mid X_{\theta_k}^{\gamma,\phi,\theta} = y\Big).
\end{aligned}$$

Therefore if ϕ is absolutely continuous, then the induction hypothesis and the time homogeneous Markov property of the pair $(X^{\gamma,\phi,\theta},\phi)$ imply

$$\begin{aligned}&\lim_{\delta\searrow 0}\limsup_{\gamma\to\infty}\gamma^{-1}\log\sup_{|x-\phi_0|<\delta}P_x\Bigl(\sup_{0\le t\le\theta_{k+1}}|X_t^{\gamma,\phi,\theta}-\phi_t|<\delta\Bigr)\\ &\qquad\le-\sum_{i=0}^{k}\int_{\theta_i}^{\theta_{i+1}}\Lambda^{\phi_{\theta_i}}(\phi_t,\dot\phi_t)\,dt.\end{aligned}\tag{6.2}$$

Otherwise, either $(\phi_t\colon 0\le t\le\theta_k)$ or $(\phi_t\colon \theta_k\le t\le\theta_{k+1})$ is not absolutely continuous, hence the left-hand side of (6.2) equals $-\infty$. This completes the induction step and establishes the lemma. □

LEMMA 6.2 (Local upper bound). *For each* $\phi\in C_{[0,T]}(R^d)$, *and* $x_0\in R^d$,

$$\lim_{\delta\searrow 0}\limsup_{\gamma\to\infty}\gamma^{-1}\log\sup_{|x-x_0|<\delta}P_x\Bigl(\sup_{0\le t\le T}|X_t^\gamma-\phi_t|<\delta\Bigr)\le-\Gamma(\phi,x_0).\tag{6.3}$$

PROOF. Without loss of generality, we may assume that ϕ is continuous and $\phi_0=x_0$, since otherwise the left-hand side of (6.3) equals $-\infty$. Fix $\varepsilon>0$, and choose $\delta>0$ such that

$$(1+\varepsilon)^{-1}\le\frac{d\nu^+(x)}{d\nu^+(x')}\le(1+\varepsilon)\quad\text{and}\quad(1+\varepsilon)^{-1}\le\frac{d\nu^-(x)}{d\nu^-(x')}\le(1+\varepsilon),$$

whenever $|x-x'|<2\delta$. Appeal to the uniform continuity of ϕ on $[0,T]$ to choose a partition $\theta=(\theta_0,\dots,\theta_{J(\theta)})$ of $[0,T]$ such that $\sup_{\theta_i\le t\le\theta_{i+1}}|\phi_t-\phi_{\theta_i}|<\delta$ for each $i\in\{0,\dots,J(\theta)-1\}$.

Let N_T^γ and $N_T^{\gamma,\phi,\theta}$ denote, respectively, the number of jumps of X^γ and $X^{\gamma,\phi,\theta}$ in the interval $[0,T]$. By the choice of θ and equation (5.3), for each $x\in R^d$, $\gamma>0$ and $B>0$,

$$\begin{aligned}&P_x\Bigl(\sup_{0\le t\le T}|X_t^\gamma-\phi_t|<\delta\Bigr)\\ &\qquad\le P_x\Bigl(\sup_{0\le t\le T}|X_t^\gamma-\phi_t|<\delta,\ N_T^\gamma<\gamma B\Bigr)+P_x(N_T^\gamma\ge\gamma B)\\ &\qquad\le e^{\gamma\varepsilon(mT+B)}P_x\Bigl(\sup_{0\le t\le T}|X_t^{\gamma,\phi,\theta}-\phi_t|<\delta,\ N_T^{\gamma,\phi,\theta}<\gamma B\Bigr)\\ &\qquad\quad+P_x(N_T^\gamma\ge\gamma B)\\ &\qquad\le e^{\gamma\varepsilon(mT+B)}P_x\Bigl(\sup_{0\le t\le T}|X_t^{\gamma,\phi,\theta}-\phi_t|<\delta\Bigr)+P_x(N_T^\gamma\ge\gamma B),\end{aligned}\tag{6.4}$$

where the second step uses the fact that $\log(1+\varepsilon)\le\varepsilon$. By hypothesis, uniformly for all initial states, N_T^γ is stochastically dominated by a Poisson random variable with mean γmT. Therefore, if ϕ is not absolutely continuous, then inequality (6.4) along with Lemma 6.1 and choice of arbitrarily large B on the right-hand side imply that the left-hand side of (6.3) equals $-\infty$, and the lemma holds.

If ϕ is absolutely continuous, then the choice of θ and Lemma 3.3 applied to the definition of $\Lambda^{\phi_{\theta_i}}$ imply

$$\begin{aligned}\sum_{i=0}^{J(\theta)-1}\int_{\theta_i}^{\theta_{i+1}}\Lambda^{\phi_{\theta_i}}(\phi_t,\dot\phi_t)\,dt &\geq \sum_{i=0}^{J(\theta)-1}\int_{\theta_i}^{\theta_{i+1}}\big((1+\varepsilon)^{-1}\Lambda^{\phi_t}(\phi_t,\dot\phi_t)+\varepsilon em\big)\,dt \\ &= (1+\varepsilon)^{-1}\Gamma(\phi,x_0)+\varepsilon emT.\end{aligned} \tag{6.5}$$

Appeal to Condition 2.1 to choose B large enough so that

$$\limsup_{\gamma\to\infty}\gamma^{-1}\log\sup_x P_x(N_T^\gamma\geq\gamma B)\leq-\Gamma(\phi,x_0).$$

Then inequality (6.4), together with Lemma 6.1, inequality (6.5), and the choice of B, implies that

$$\begin{aligned}&\lim_{\delta\searrow 0}\limsup_{\gamma\to\infty}\gamma^{-1}\log\sup_{|x-x_0|<\delta}P_x\Big(\sup_{0\leq t\leq T}|X_t^\gamma-\phi_t|<\delta\Big)\\ &\qquad\leq\big(-(1+\varepsilon)^{-1}\Gamma(\phi,x_0)-\varepsilon emT+\varepsilon(mT+B)\big)\bigvee(-\Gamma(\phi,x_0)).\end{aligned}$$

The lemma follows by the arbitrariness of $\varepsilon>0$. $\square$

LEMMA 6.3 (Exponential tightness). *Let $(x^\gamma\colon\gamma>0)$ be a sequence such that $\lim_{\gamma\to\infty}x^\gamma=x_0$. For each $\alpha>0$ there exists a compact $K_\alpha\subset C_{[0,T]}(R^d)$ such that*

$$\limsup_{\gamma\to\infty}\gamma^{-1}\log P_{x^\gamma}\big((\tilde X_t^\gamma\colon 0\leq t\leq T)\notin K_\alpha\big)\leq-\alpha.$$

PROOF. The lemma follows by Lemma 5.58 of Shwartz and Weiss (1995) through a straightforward adaptation of their Corollary 5.8 so as to incorporate the continuous variaton of the measures $\nu(x)$, $x\in R^d$. $\square$

LEMMA 6.4 (Upper bound). *For any Borel measurable $S\subset C_{[0,T]}(R^d)$, $x_0\in R^d$ and sequence $(x^\gamma\colon\gamma>0)$ such that $\lim_{\gamma\to\infty}x^\gamma=x_0$,*

$$\limsup_{\gamma\to\infty}\gamma^{-1}\log P_{x^\gamma}\big((\tilde X_t^\gamma\colon 0\leq t\leq T)\in S\big)\leq-\inf_{\phi\in\overline{S}}\Gamma(\phi,x_0).$$

PROOF. Fix $\varepsilon>0$. For each $\phi\in\overline{S}$ appeal to Lemma 6.2 and Lemma 3.2 to choose a $\delta_\phi>0$ such that

$$\limsup_{\gamma\to\infty}\gamma^{-1}\log P_{x^\gamma}\big((\tilde X_t^\gamma\colon 0\leq t\leq T)\in B(\phi,\delta_\phi)\big)\leq-\Gamma(\phi,x_0)+\varepsilon,$$

and appeal to Lemma 6.3 to choose a compact subset K of $C_{[0,T]}(R^d)$ such that

$$\limsup_{\gamma\to\infty}\gamma^{-1}\log P_{x^\gamma}\big((\tilde X_t^\gamma\colon 0\leq t\leq T)\notin K\big)\leq-\Big(\frac{1}{\varepsilon}\wedge\inf_{\phi\in\overline{S}}\Gamma(\phi,x_0)\Big).$$

By the compactness of $\overline{S}\cap K$ there exists a finite subset $\{\phi^1, \dots, \phi^I\} \subset \overline{S}$ such that $\overline{S} \cap K \subset \bigcup_{i=1}^{I} B(\phi^i, \delta_{\phi^i})$, hence for each $\gamma > 0$,

$$P_{x^\gamma}\big((\tilde{X}_t^\gamma\colon 0 \le t \le T) \in S\big) \le P_{x^\gamma}\big((\tilde{X}_t^\gamma\colon 0 \le t \le T) \notin K\big) + \sum_{i=1}^{I} P_{x^\gamma}\big((\tilde{X}_t^\gamma\colon 0 \le t \le T) \in B(\phi^i, \delta_{\phi^i})\big).$$

This in turn implies

$$\begin{aligned}
&\limsup_{\gamma\to\infty} \gamma^{-1} \log P_{x^\gamma}\big((\tilde{X}_t^\gamma\colon 0 \le t \le T) \in S\big)\\
&\quad\le \left(-\left(\frac{1}{\varepsilon} \wedge \inf_{\phi\in\overline{S}} \Gamma(\phi, x_0)\right)\right) \bigvee \max_{1\le i\le I}\{-\Gamma(\phi^i, x_0) + \varepsilon\}\\
&\quad\le \left(-\left(\frac{1}{\varepsilon} \wedge \inf_{\phi\in\overline{S}} \Gamma(\phi, x_0)\right)\right) \bigvee \Big(-\inf_{\phi\in\overline{S}} \Gamma(\phi, x_0) + \varepsilon\Big).
\end{aligned}$$

The lemma now follows by the arbitrariness of $\varepsilon > 0$. □

7. Goodness of the rate function. This section concludes the proof of Theorem 2.1 by establishing the goodness of the rate function $\Gamma(\cdot, x_o)$.

LEMMA 7.1 (Lower semicontinuity). *Given $x_0 \in R^d$, the function $\Gamma(\cdot, x_0)\colon C_{[0,T]}(R^d) \to R_+ \cup \{+\infty\}$ is lower semicontinuous.*

PROOF. Let $(\phi^m\colon m > 0)$ be a sequence such that $\phi^m \to \phi$ in $C_{[0,T]}(R^d)$. To prove the lemma, it is enough to show that $\Gamma(\phi, x_0) \le \liminf_{m\to\infty} \Gamma(\phi^m, x_0)$. Fix $\varepsilon > 0$ and a sequence $x^\gamma \to x_0$. By Lemma 6.2 there exists a $\delta > 0$ such that

$$\limsup_{\gamma\to\infty} \gamma^{-1} \log P_{x^\gamma}\big((\tilde{X}_t^\gamma\colon 0 \le t \le T) \in B(\phi, \delta)\big) \le -\Gamma(\phi, x_0) + \varepsilon. \tag{7.1}$$

Let M_ε be such that $\sup_{0\le t\le T} |\phi_t^m - \phi_t| < \delta$ whenever $m > M_\varepsilon$. By Lemma 5.3,

$$\begin{aligned}
&\liminf_{\gamma\to\infty} \gamma^{-1} \log P_{x^\gamma}\big((\tilde{X}_t^\gamma\colon 0 \le t \le T) \in B(\phi, \delta)\big)\\
&\quad\ge -\inf_{\phi'\in B(\phi,\delta)} \Gamma(\phi', x_0)\\
&\quad\ge -\Gamma(\phi^m, x_0)
\end{aligned} \tag{7.2}$$

for all $m > M_\varepsilon$. Inequalities (7.1) and (7.2) imply that $\Gamma(\phi^m, x_0) \ge \Gamma(\phi, x_0) - \varepsilon$ whenever $m > M_\varepsilon$, and the lemma is proved. □

LEMMA 7.2 (Goodness). *The rate function $\Gamma(\cdot, x_o)$ is good.*

PROOF. In view of Lemma 1.2.18.b of Dembo and Zeitouni (1992), the lemma is implied by Lemmas 5.3, 6.3 and 7.1. □

REFERENCES

ALANYALI, M. and HAJEK, B. (1998). On large deviations in load sharing networks. *Ann. Appl. Probab.* **8** 67–97.

AZENCOTT, R. and RUGET, G. (1977). Melanges d'équations differentialles et grands écarts à la loi des grand nombres. *Z. Wahrsch. Verw. Gebiete* **38** 1–54.

BLINOVSKII, V. M. and DOBRUSHIN, R. L. (1994). Process level large deviations for a class of piecewise homogeneous random walks. In *The Dynkin Festschrift: Markov Processes and Their Applications* 1–59. Birkhäuser, Boston.

DEMBO, A. and ZEITOUNI, O. (1992). *Large Deviations Techniques and Applications*. Jones and Bartlett, Boston.

DINWOODIE, I. H. and ZABELL, S. L. (1992). Large deviations for exchangeable random vectors. *Ann. Probab.* **20** 1147–1166.

DUPUIS, P. and ELLIS, R. S. (1992). Large deviations for Markov processes with discontinuous statistics II. *Probab. Theory Related Fields* **91** 153–194.

DUPUIS, P. and ELLIS, R. S. (1995). The large deviation principle for a general class of queueing systems I. *Trans. Amer. Math. Soc.* **347** 2689–2751.

DUPUIS, P. and ELLIS, R. S. (1996). *A Weak Convergence Approach to the Theory of Large Deviations*. Wiley, New York.

DUPUIS, P., ELLIS, R. S. and WEISS, A. (1991). Large deviations for Markov processes with discontinuous statistics I: general upper bounds. *Ann. Probab.* **19** 1280–1297.

IGNATYUK, I. A., MALYSHEV, V. and SCHERBAKOV, V. V. (1994). Boundary effects in large deviation problems. *Russian Math. Surveys* **49** 41–99.

NAGOT, I. (1995). Grandes déviations pour les processus d'apprentissage lent à statistiques discontinues sur une surface. Thèse de Docteur en Sciences, Université Paris XI Orsay, U.F.R. Scientifique d'Orsay.

ROCKAFELLAR, R. T. (1970). *Convex Analysis*. Princeton Univ. Press.

SHWARTZ, A. and WEISS, A. (1995). *Large Deviations for Performance Analysis, Queues, Communication and Computing*. Chapman & Hall, London.

BELL LABORATORIES
LUCENT TECHNOLOGIES
101 CRAWFORDS CORNER ROAD
HOLMDEL, NEW JERSEY 07733-3030
E-MAIL: murat@dnrc.bell-labs.com

DEPARTMENT OF ELECTRICAL
AND COMPUTER ENGINEERING
AND COORDINATED SCIENCE LABORATORY
UNIVERSITY OF ILLINOIS
1308 WEST MAIN STREET
URBANA, ILLINOIS 61801
E-MAIL: b-hajek@uiuc.edu