Probab. Theory Relat. Fields 91, 153–194 (1992)

Probability Theory and Related Fields

Large deviations for Markov processes with discontinuous statistics, II: random walks

Paul Dupuis* **and Richard S. Ellis****

Department of Mathematics and Statistics, University of Massachusetts, Amherst, M/A 01003, USA

Received July 5, 1990, in revised form May 25, 1991

Summary. Let μ^1 and μ^2 be Borel probability measures on $\mathbb{R}^d$ with finite moment generating functions. The main theorem in this paper proves the large deviation principle for a random walk whose transition mechanism is governed by μ^1 when the walk is in the left halfspace $\Lambda^1 = \{x \in \mathbb{R}^d : x_1 \leqq 0\}$ and whose transition mechanism is governed by μ^2 when the walk is in the right halfspace $\Lambda^2 = \{x \in \mathbb{R}^d : x_1 > 0\}$. When the measures μ^1 and μ^2 are equal, the main theorem reduces to Cramér's Theorem.

1 Introduction

An extensive theory has been developed for analyzing large deviation phenomena of d-dimensional Markov processes having generators $\mathscr{L}$ with components that depend smoothly upon the spatial parameter $x \in \mathbb{R}^d$. Three basic examples are diffusion processes, continuous-time jump Markov processes, and Markov chains. In the first case, the generator has the form

$$\mathscr{L}f(x) = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(x)\frac{\partial^2 f(x)}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i(x)\frac{\partial f(x)}{\partial x_i},$$

in which the diffusion matrix $a(x) = \{a_{ij}(x), i, j = 1, \ldots, n\}$ and the drift vector $b(x) = \{b_i(x), i = 1, \ldots, n\}$ are smooth functions of x. A typical

* This research was supported in part by a grant from the National Science Foundation (NSF-DMS-8902333)

** This research was supported in part by a grant from the National Science Foundation (NSF-DMS-8901138) and in part by a Lady Davis Fellowship while visiting the Faculty of Industrial Engineering and Management at the Technion during the spring semester of 1989

generator that arises in the second case is

$$\mathscr{L}f(x) = \int_{\mathbb{R}^d\setminus\{0\}} [f(x+v) - f(x)]\mu_x(dv),$$

in which the jump measures $\{\mu_x\}$ depend smoothly upon x. In the third case, the generator has the form

$$\mathscr{L}f(x) = \int_{\mathbb{R}^d} [f(y) - f(x)]\pi(x, dy),$$

in which the transition probability $\pi(x, dy)$ satisfies the Feller property.

In a variety of applications, however, Markov processes arise naturally for which the smooth dependence of the components of the generator $\mathscr{L}$ upon x is violated. In such a case, we speak of a Markov process with "discontinuous statistics."

One application of Markov processes with discontinuous statistics is to communication channels incorporating a "hard limiter" in a phase-locked loop, which is a form of a suboptimal nonlinear filter [25]. Such a communication channel may be modeled by a diffusion process with a smooth diffusion matrix $a(x)$ and with a drift vector $b(x)$ that changes discontinuously as x crosses a smooth boundary in $\mathbb{R}^d$. The large deviation principle for a restricted class of diffusion processes with discontinuous drifts satisfying a certain stability condition is considered by Korostelev and Leonov [22], [23]. They use continuous mapping techniques that may be found in the literature.

A second application of Markov processes with discontinuous statistics is to queueing networks consisting of d queues or consisting of a single queue with d classes of customers [27]. Such a network may be modeled by a continuous-time jump Markov process $\{X(t), t \geqq 0\}$ with state space the nonnegative orthant of $\mathbb{R}^d$. For $i \in \{1, \ldots, d\}$, the i'th coordinate of $X(t)$ denotes the length of the i'th queue at time t or the number of customers of class i awaiting service at time t. In general, the queueing network has one set of statistics when all the queues are nonempty or when all the classes of customers are present. However, in many cases the behavior of the queueing network changes abruptly and discontinuously when one or more of the queues become empty or when one or more classes of customers disappear. We may model such a network by stipulating that $X(t)$ have one set of jump rates and directions in the positive orthant of $\mathbb{R}^d$, but that as $X(t)$ moves to one of the coordinate hyperplanes, the jump rates and directions change discontinuously. Since $X(t)$ may trivially be extended to a process with state space all of $\mathbb{R}^d$, we see that the queueing network may be modeled by a continuous-time jump Markov process with jump measures $\{\mu_x\}$ that change discontinuously as x crosses one of the coordinate hyperplanes. Large deviation phenomena for a special class of queueing networks are studied in the paper by Dupuis et al. [13].

A third application of Markov processes with discontinuous statistics is to random motion in discontinuous media. Such motion may be modeled by a Markov chain with state space $\mathbb{R}^d$ partitioned into finitely many sets $\{\Lambda^i, i = 1, \ldots, N\}$, in each of which the Markov chain has a different smooth transition mechanism. The transition mechanisms change abruptly and discontinuously when the chain crosses the boundaries separating the sets $\{\Lambda^i\}$. The present paper is devoted to the large deviation analysis of such a Markov chain with discontinuous statistics in the simplest case $N = 2$. It is defined in (1.1) below. This Markov chain is a random walk that generalizes a random walk based on sums of i.i.d. random vectors, about which there is a voluminous literature.

The large deviation principle for the Markov chain (1.1) suitably normalized is proved in our main theorem, Theorem 2.1. This theorem generalizes Cramér's Theorem, which gives the large deviation principle for random walks based on sums of i.i.d. random vectors. Cramér's Theorem is stated in Theorem 1.2 below.

We first give a basic definition.

Definition 1.1 Let $\{Q_n, n \in \mathbb{N}\}$ be a sequence of Borel probability measures on $\mathbb{R}^d$, some $d \in \mathbb{N}$, and let L be an extended real-valued function mapping $\mathbb{R}^d$ into $[0, \infty]$. We say that $\{Q_n\}$ satisfies the large deviation principle with rate function L if the following three conditions hold.

(a) *Compact level sets.* For each $M < \infty$ the set $\{\beta \in \mathbb{R}^d : L(\beta) \leqq M\}$ is compact.

(b) *Upper large deviation bound.* For each closed set F in $\mathbb{R}^d$

$$\limsup_{n\to\infty} n^{-1} \log Q_n\{F\} \leqq -\inf_{\beta\in F} L(\beta) .$$

(c) *Lower large deviation bound.* For each open set G in $\mathbb{R}^d$

$$\liminf_{n\to\infty} n^{-1} \log Q_n\{G\} \geqq -\inf_{\beta\in G} L(\beta) .$$

The Markov chain that is the subject of the present paper is easy to define. Let μ^1 and μ^2 be Borel probability measures on $\mathbb{R}^d$ such that the cumulant generating functions

$$H^i(\alpha) \doteq \log \int_{\mathbb{R}^d} \exp\langle\alpha, y\rangle \mu^i(\mathrm{d}y), \quad i = 1, 2 ,$$

are both finite for all $\alpha \in \mathbb{R}^d$. We consider a Markov chain whose transition mechanism is governed by μ^1 when the walk is in the left halfspace $\Lambda^1 = \{x \in \mathbb{R}^d : x_1 \leqq 0\}$ and whose transition mechanism is governed by μ^2 when the walk is in the right halfspace $\Lambda^2 = \{x \in \mathbb{R}^d : x_1 > 0\}$. The discontinuity occurs across the boundary $\{x \in \mathbb{R}^d : x_1 = 0\}$. Specifically, let $\{X_n^i, n \in \mathbb{N}, i = 1, 2\}$ be a set of independent random vectors with probability distributions $P\{X_n^i \in \mathrm{d}x\} = \mu^i(\mathrm{d}x)$. We consider the stochastic process $\{S_n, n \in \mathbb{N}\}$, where $S_0 = 0$ and S_n is defined recursively by the formula

$$S_{n+1} = S_n + 1_{\{S_n\in\Lambda^1\}} X_{n+1}^1 + 1_{\{S_n\in\Lambda^2\}} X_{n+1}^2 . \tag{1.1}$$

For $i = 1, 2$, $1_{\{S_n\in\Lambda^i\}}$ denotes the indicator function of the set $\{S_n \in \Lambda^i\}$.

Our goal is to prove the large deviation principle for $\{S_n/n\}$. It is stated in Theorem 2.1 and the rate function is given in formulas (2.1)–(2.4). The stochastic process $\{S_n\}$ has an obvious asymmetry in that we have (arbitrarily) included the boundary $\{x \in \mathbb{R}^d : x_1 = 0\}$ in the halfspace Λ^1. If the interiors of the two halfspaces Λ^1 and Λ^2 communicate in a manner that is formalized in Hypothesis (H), stated just below, then the definition of the transition mechanism on the boundary does not affect the large deviation principle for $\{S_n/n\}$.

In Theorem 2.1, we impose the following weak hypothesis on the supports, supp μ^1 and supp μ^2, of the probability measures μ^1 and μ^2. The hypothesis is needed only in the proof of the lower large deviation bound, not in the proof of the upper large deviation bound or in the proof of compact level sets.

Hypothesis (H). $\operatorname{supp} \mu^i \cap \operatorname{int} \Lambda^{3-i} \neq \emptyset \qquad \text{for} \quad i = 1, 2.$

This hypothesis is natural in that it allows the interiors of the two half-spaces Λ^1 and Λ^2 to communicate. We emphasize that only under Hypothesis (H) is the unique feature of the process $\{S_n\}$ present; namely, the existence of two separate transition mechanisms, each of which has a positive probability of being activated for arbitrarily large times n. For example, if Hypothesis (H) fails for $i = 1$, then since the random walk starts at the origin, which is a point in the halfspace Λ^1, the random walk will never enter the halfspace Λ^2 and so the transition mechanism corresponding to μ^2 will never be activated. Even when Hypothesis (H) fails, the large deviation principle for $\{S_n/n\}$ is valid, but the rate function is no longer given by formulas (2.1)–(2.4).

The process $\{S_n, n \in \mathbb{N}\}$ defined in (1.1) is a Markov chain with state space $\mathbb{R}^d$ and transition probability

$$(1.2) \qquad \pi(x, \mathrm{d}y) = 1_{\{x \in \Lambda^1\}} \mu^1(\mathrm{d}y - x) + 1_{\{x \in \Lambda^2\}} \mu^2(\mathrm{d}y - x) .$$

Since μ^1 and μ^2 are arbitrary measures on $\mathbb{R}^d$ that have finite cumulant generating functions, satisfy Hypothesis (H), but have no imposed relationships between them, the transition probability $\pi(x, \mathrm{d}y)$ satisfies none of the usual absolute continuity, irreducibility, or smoothness conditions assumed in other large deviation studies. As a consequence, one is not able to analyze the process (1.1) by any of the usual methods. In fact, a number of completely new techniques are required.

The present paper is the second in a series of papers that treat large deviation phenomena for Markov processes with discontinuous statistics. The first paper by Dupuis et al. [12] proved a process-level upper large deviation bound for a general class of Markov processes with generators that contain both a diffusion piece and a jump piece. This paper also proved a process-level upper large deviation bound for continuous-time Markov process obtained by linear interpolation from random walks that model motion in "totally discontinuous" media. These random walks include as a special case the Markov chain analyzed in the present paper.

The process-level upper large deviation bounds proved in Dupuis et al. [12] apply to a general class of Markov processes with discontinuous statistics. However, in order the obtain the full large deviation principle (upper bound and lower bound), some sacrifice in generality is needed. It is our judgment that if one hopes to understand large deviation phenomena in the case of continuous-time Markov processes with discontinuous statistics, then large deviation phenomena for discrete-time Markovian models, such as the one treated in the present paper, must be understood first. In future work, we plan to apply the insights gained in the present paper to analyze continuous-time Markov processes with discontinuous statistics, including diffusion processes that model communication channels with a hard limiter in a phase-locked loop and jump Markov processes that model queueing networks. The methods of Korostelev and Leonov [22], [23], mentioned at the beginning of this introduction, are restricted to diffusions for which a certain mapping of processes is continuous. In such cases, standard methods involving the contraction principle may be used. Our main interest, and the source of much of the difficulty, is to deal with the general case, where such continuity does not hold. In the applications both to diffusion processes that model communication channels with a hard limiter in a phase-lock loop and to jump Markov processes that model queueing networks, no such continuous mapping of processes exists.

Cramér's Theorem, which treats random walks on $\mathbb{R}^d$ based on sums of i.i.d. random vectors, is one of the basic results in the theory of large deviations. The following theorem is proved, for example, in [15]. In its original form, the theorem goes back to Cramér [3].

Theorem 1.2 (Cramér) *Let μ be a Borel probability measure on $\mathbb{R}^d$ such that the cumulant generating function*

$$H(\alpha) \doteq \log \int_{\mathbb{R}^d} \exp\langle \alpha, y\rangle \mu(\mathrm{d}y)$$

is finite for all $\alpha \in \mathbb{R}^d$. Let $\{X_j, j \in \mathbb{N}\}$, be a sequence of i.i.d. random vectors with probability distribution μ. For $n \in \mathbb{N}$, define $S_n = \sum_{j=1}^n X_j$ and let Q_n denote the probability distribution of $n^{-1}S_n$. The following conclusions hold.

(a) *$\{Q_n\}$ satisfies the large deviation principle with rate function L given by the Legendre-Fenchel transform of H:*

$$L(\beta) = \sup_{\alpha \in \mathbb{R}^d} \{\langle \alpha, \beta\rangle - H(\alpha)\} \qquad \text{for } \beta \in \mathbb{R}^d . \tag{1.3}$$

(b) *The function $L(\beta)$ is lower semicontinuous, superlinear, and essentially strictly convex. In addition, $L(\beta) > L(\bar{\beta}) = 0$ for all $\beta \neq \bar{\beta}$, where $\bar{\beta}$ is the mean $\int_{\mathbb{R}^d} y\mu(\mathrm{d}y)$. If the convex hull of the support of μ is all of $\mathbb{R}^d$, then $L(\beta)$ is also real analytic and strictly convex on $\mathbb{R}^d$.*

Many generalizations of Cramér's Theorem to Markov processes have been obtained. A number of authors, including de Acosta [4], Ellis [14], Freidlin and Wentzell [18], Gärtner [19], and Iscoe et al. [20], extended the techniques used in the proof of Cramér's Theorem to prove the large deviation principle for sequences of random vectors, including partial sums of Markov chains, for which a suitably smooth limiting cumulant generating function exists.

The extensions found by these authors use methods of convex analysis. In a series of papers beginning in 1975, Donsker and Varadhan [8–11] proved the large deviation principle for the empirical measure (level – 2) and the empirical process (level – 3) of Markov processes and Markov chains satisfying certain absolute continuity, irreducibility, and smoothness conditions. They obtained as a corollary an extension of Cramér's Theorem to i.i.d. random vectors taking values in a Banach space. Results related to the work of Donsker and Varadhan and generalizations were found by a number of authors, including Bolthausen [2], de Acosta [4–6], Deuschel and Stroock [7], Ellis [16], Ellis and Wyner [17], Jain [21], and Stroock [29], some of whom used methods of infinite dimensional convex analysis. Another important direction was pursued by Azencott and Ruget [1], Freidlin and Wentzell [18], and Wentzell [30–33], who proved the process-level large deviation principle for Markov processes with continuous statistics.

Unfortunately, none of the techniques used or the results obtained by any of the above authors are applicable to the analysis of the Markov chain (1.1).

In the next section, we state the main theorem, Theorem 2.1. In Sect. 3, we interpret the rate function in Theorem 2.1 and give examples of rate functions which – in contrast of the real analytic, strictly convex rate functions in Cramér's Theorem and its many generalizations – are not everywhere differentiable, are convex but not strictly convex, and are not convex at all. In Sect. 4, we collect facts about Legendre-Fenchel transforms needed in the rest of the paper and prove that the function L appearing in Theorem 2.1 has compact level sets. Finally, in Sect. 5, we prove the upper and lower large deviation bounds in Theorem 2.1.

2 Statement of the main theorem

The Markov chain $\{S_n, n \in \mathbb{N}\}$ that is the subject of the present paper is defined in (1.1). Theorem 2.1 is a new large deviation result for $\{S_n/n\}$. This theorem assumes Hypothesis (H), stated in Sect. 1 and again below.

Recall that μ^1 and μ^2 are Borel probability measures on $\mathbb{R}^d$ such that the cumulant generating functions

$$H^i(\alpha) \doteq \log \int_{\mathbb{R}^d} \exp\langle \alpha, y\rangle \mu^i(dy), \qquad i = 1, 2,$$

are both finite for all $\alpha \in \mathbb{R}^d$. Recall also Hypothesis (H) in Sect. 1, which states that the supports of μ^1 and μ^2 satisfy

$$\operatorname{supp} \mu^i \cap \operatorname{int} \Lambda^{3-i} \neq \emptyset \qquad \text{for } i = 1, 2,$$

where Λ^1 denotes the closed left halfspace $\{x \in \mathbb{R}^d : x_1 \leqq 0\}$ and Λ^2 denotes the open right halfspace $\{x \in \mathbb{R}^d : x_1 > 0\}$.

The rate function in Theorem 2.1 is now defined in several pieces. If $\beta \in \mathbb{R}^d$ lies along the boundary $\Lambda^0 = \{x \in \mathbb{R}^d : x_1 = 0\}$ (i.e., if the one-component

$\beta_1 = 0$), then we define

$$\tilde{L}^0(\beta) = \inf\{\varrho^1 L^1(\beta^1) + \varrho^2 L^2(\beta^2) : \varrho^1 \geqq 0,\ \varrho^2 \geqq 0,\ \varrho^1 + \varrho^2 = 1,\ \beta^1 \in \mathbb{R}^d, \beta^2 \in \mathbb{R}^d,\ (\beta^1)_1 \geqq 0,\ (\beta^2)_1 \leqq 0,\ \varrho^1\beta^1 + \varrho^2\beta^2 = \beta\}, \tag{2.1}$$

where $(\beta^i)_1, i = 1, 2$, denotes the one-component of β^i and $L^i, i = 1, 2$, denotes the Legendre-Fenchel transform of H^i:

$$L^i(\beta) = \sup_{\alpha \in \mathbb{R}^d} \{\langle \alpha, \beta \rangle - H^i(\alpha)\} \qquad \text{for } \beta \in \mathbb{R}^d . \tag{2.2}$$

We also define for $i = 1, 2$ and $\beta \in \mathbb{R}^d$

$$\tilde{L}^i(\beta) = \inf\{\gamma^0 \tilde{L}^0(\beta^0) + \gamma^i L^i(\beta^i) : \gamma^0 \geqq 0,\ \gamma^i \geqq 0,\ \gamma^0 + \gamma^i = 1,\ \beta^0 \in \mathbb{R}^d,\ \beta^i \in \mathbb{R}^d,\ (\beta^0)_1 = 0,\ \gamma^0\beta^0 + \gamma^i\beta^i = \beta\} . \tag{2.3}$$

Finally, we define

$$L(\beta) = \begin{cases} \tilde{L}^0(\beta) & \text{if } \beta_1 = 0 \\ \tilde{L}^1(\beta) & \text{if } \beta_1 < 0 \\ \tilde{L}^2(\beta) & \text{if } \beta_1 > 0 . \end{cases} \tag{2.4}$$

We now state our main theorem.

Theorem 2.1. *Let μ^1 and μ^2 be Borel probability measures on $\mathbb{R}^d$ that have finite cumulant generating functions and that satisfy Hypothesis (H). For $n \in \mathbb{N}$ let Q_n denote the probability distribution of S_n/n, where S_n is defined recursively in* (1.1). *The following conclusions hold.*

(a) *$\{Q_n\}$ satisfies the large deviation principle with rate function L defined in* (2.1) – (2.4).

(b) *The function $L(\beta)$ is lower semicontinuous and superlinear and is convex on each halfspace $\{\beta \in \mathbb{R}^d : (-1)^i\beta_1 \geqq 0\}, i = 1, 2$. However, $L(\beta)$ is in general not convex on $\mathbb{R}^d$.*

In Proposition 4.3 in Sect. 4, we prove that the function L has compact level sets and that L has the other properties stated in part (b) of Theorem 2.1. The large deviation bounds for $\{Q_n\}$ with rate function L are proved in Sect. 5.

Although the rate function L in Theorem 2.1 looks complicated, it has a natural interpretation in terms of which vectors the Markov chain "tracks" and the asymptotic fractions of time that the Markov chain spends in the halfspaces Λ^1 and Λ^2 when the chain is conditioned on a suitable sequence of events. This interpretation is given in Sect. 3. The constraints $(\beta^1)_1 \geqq 0$, $(\beta^2)_1 \leqq 0$ in (2.1) appearing in the definition of the rate function L have natural interpretations both in terms of the upper large deviation bound (identification of those velocities that can be "tracked" with positive probability [Lemma 5.2]) and in terms of the lower large deviation bound (a guarantee that a process arising from S_n by a certain change of measure is stable [proof of Lemma 5.10]). If the constraints were absent from (2.1), then in general the resulting function would no longer be the rate function for S_n/n.

We close this section by pointing out the limitations of the convex analysis approach to large deviations [4], [14], [18], [19]. We consider the limiting cumulant generating function of $\{S_n\}$, which for $\alpha \in \mathbb{R}^d$ is defined by the formula

$$H(\alpha) = \lim_{n \to \infty} n^{-1} \log E\{\exp[\langle \alpha, S_n \rangle]\} .$$

Theorem 2.1 and a result of Varadhan (see Sect. II.7 of [15]) imply that

$$H(\alpha) = \sup_{\beta \in \mathbb{R}^d} \{\langle \alpha, \beta \rangle - L(\beta)\} . \tag{2.5}$$

According to the convex analysis approach, the upper large deviation bound holds for the distributions $\{Q_n\}$ of $\{S_n/n\}$ with the upper rate function $I(\beta) = \sup_{\alpha \in \mathbb{R}^d} \{\langle \alpha, \beta \rangle - H(\alpha)\}$, which is automatically convex. On the other hand, (2.5) implies that I equals the largest lower semicontinuous, convex function majorized by L. We conclude that whenever L is not convex (see examples in the next section), the function I cannot be the rate function for $\{Q_n\}$ and that for some closed set F in $\mathbb{R}^d$ the upper large deviation bound in terms of I cannot be so tight as the upper large deviation bound in terms of L.

Notation. Throughout this paper, except in a small number of instances, a superscript $i \in \{0, 1, 2\}$ is used to identify quantities relative to the respective sets $\{\beta \in \mathbb{R}^d : \beta_1 = 0\}$, the halfspace $\{\beta \in \mathbb{R}^d : \beta_1 \leqq 0\}$ (or its interior), and the halfspace $\{\beta \in \mathbb{R}^d : \beta_1 > 0\}$ (or its closure). Thus, we have measures μ^1 and μ^2; vectors β^0, β^1, and β^2; and functions $\tilde{L}^0$, $\tilde{L}^1$ and $\tilde{L}^2$. In the small number of instances where a superscript denotes a power, the quantity raised to the power is enclosed in square brackets. Thus, $[x]^2$ denotes x-squared.

3 Discussion of the main theorem

Before examining Theorem 2.1, we first interpret the rate function in Cramér's Theorem, Theorem 1.2. In order to simplify the discussion, we assume that the support of the measure μ is all of $\mathbb{R}^d$, which implies that the Legendre-Fenchel transform $L(\beta)$ in (1.3) is finite for all $\beta \in \mathbb{R}^d$ [Theorem 1.2 b]. For all $\beta \in \mathbb{R}^d$ the large deviation principle gives the limit

$$\lim_{\varepsilon \to 0} \lim_{n \to \infty} n^{-1} \log P\{S_n/n \in B(\beta, \varepsilon)\} = -L(\beta) . \tag{3.1}$$

Let $\bar{\beta}$ equal the mean $\int_{\mathbb{R}^d} y\mu(dy)$. The law of large numbers implies that for each $\varepsilon > 0$ $P\{S_n/n \in B(\bar{\beta}, \varepsilon)\} \to 1$ as $n \to \infty$; i.e., as $n \to \infty$ the random walk tracks $\bar{\beta}$ with probability approaching 1. This, of course, is consistent with (3.1) and the fact that $L(\bar{\beta}) = 0$. For all $\beta \neq \bar{\beta}$, we express (3.1) by the formula

$$P\{S_n/n \in B(\beta, \varepsilon)\} \approx \exp[-nL(\beta)]$$

and interpret $L(\beta)$ as the "positive cost" associated with the atypical event that the random walk S_n, suitably normalized, tracks β.

We now turn to Theorem 2.1, assuming for the rest of this section that the supports of both measures μ^1 and μ^2 are all of $\mathbb{R}^d$. Then both Legendre-Fenchel transforms L^1 and L^2 are finite on all of $\mathbb{R}^d$, as are the functions $\tilde{L}^0, \tilde{L}^1, \tilde{L}^2$, and L in (2.1), (2.3), and (2.4). As in (3.1), the limit

$$\lim_{\varepsilon \to 0} \lim_{n \to \infty} n^{-1} \log P\{S_n/n \in B(\beta, \varepsilon)\} = -L(\beta) \tag{3.2}$$

is valid for all $\beta \in \mathbb{R}^d$. Let us consider the case where the one-component β_1 of β equals 0, so that $L(\beta)$ equals $\tilde{L}^0(\beta)$. We interpret $\tilde{L}^0(\beta)$, whenever it is positive, as the positive cost associated with the atypical event that the random walk S_n, suitably normalized, tracks β. If the infimum in (2.1) is attained at $\tilde{\varrho}^1, \tilde{\varrho}^2, \tilde{\beta}^1, \tilde{\beta}^2$, then the most likely way for the random walk to track β is for the random walk to track $\tilde{\beta}^1$ in the left halfspace and $\tilde{\beta}^2$ in the right halfspace; i.e. for $i = 1, 2$

$$\frac{1}{\sum_{j=1}^n 1_{\{S_{j-1} \in \Lambda^i\}}} \cdot \sum_{j=1}^n 1_{\{S_{j-1} \in \Lambda^i\}} X_j^i \to \beta^i \qquad \text{in probability as } n \to \infty .$$

When the random walk is conditioned to track β, $\tilde{\varrho}^1$ and $\tilde{\varrho}^2$ are the asymptotic fractions of time spent in the respective halfspace. The functions $\tilde{L}^i$ defined in (2.3) may be interpreted similarly.

We now consider the calculation of rate functions L in Theorem 2.1 in the case $d = 1$. When $d = 1$, formula (2.3) for $\tilde{L}^i$ takes the form

$$\tilde{L}^i(\beta) = \inf\{(1-\gamma)\tilde{L}^0(0) + \gamma L^i(\beta/\gamma) : 0 < \gamma \leqq 1\} \qquad \text{for } i = 1, 2 . \tag{3.3}$$

According to Lemma 3.1 below, the infimum in (3.3) is always attained. If it is attained at $\tilde{\gamma} \in (0, 1)$, then in tracking β the random walk, with high probability, spends at the origin an asymptotically positive fraction $(1-\tilde{\gamma})$ of the total time n, paying a cost $(1-\tilde{\gamma})\tilde{L}^0(0)$ before entering the halfline where β lies. If the infimum in (3.3) is attained at $\gamma = 1$, then in tracking β the random walk, with high probability, spends at the origin an asymptotically zero function of the total time n before entering the halfline where β lies.

We give a lemma that is useful in determining where the infimum in (3.3) is attained.

Lemma 3.1 *Let $d = 1$ and assume that the supports of both measures μ^1 and μ^2 are all of $\mathbb{R}$. The following conclusions hold.*

(a) *For $i = 1, 2,$ and $\beta \in \mathbb{R}$ satisfying $(-1)^i \beta > 0$, the function*

$$\lambda_\beta^i(\gamma) = (1-\gamma)\tilde{L}^0(0) + \gamma L^i(\beta/\gamma) \tag{3.4}$$

appearing in (3.3) *is a strictly convex function of* $\gamma \in (0, 1]$. *The infimum in* (3.3) *is always attained.*

(b) *For* $i = 1, 2$ *the infimum in* (3.3) *is attained at a (unique) point* $\tilde{\gamma} \in (0, 1)$ *if and only if for some point* $b \in \mathbb{R}$ *satisfying* $(-1)^i b > 0$

$$-H^i((L^i)'(b)) = \tilde{L}^0(0) . \tag{3.5}$$

In this case $b = \beta/\tilde{\gamma}$.

Proof. (a) For $i = 1, 2$, for $\beta \in \mathbb{R}$ satisfying $(-1)^i \beta > 0$, and for $\gamma \in (0, 1)$, we have $(L^i)''(\beta/\gamma) > 0$. Hence

$$(\lambda^i_\beta)''(\gamma) = ([\beta]^2/[\gamma]^3) \cdot (L^i)''(\beta/\gamma) > 0 .$$

Thus, $\lambda^i_\beta(\gamma)$ is a strictly convex function of $\gamma \in (0, 1)$. Since L^i is superlinear [Lemma 4.1], $\lambda^i_\beta(\gamma) \to \infty$ as $\gamma \to 0^+$ and so $\lambda^i_\beta(\gamma)$ always attains its minimum on $(0, 1]$.

(b) Since $\lambda^i_\beta(\gamma)$ is strictly convex and $\lambda^i_\beta(\gamma) \to \infty$ as $\gamma \to 0^+$, either $\lambda^i_\beta(\gamma)$ is monotonically decreasing on (0,1) or $(\lambda^i_\beta)'(\tilde{\gamma}) = 0$ for some unique point $\tilde{\gamma} \in (0, 1)$. In the first case, the infimum in (3.3) is attained at $\gamma = 1$. In the second case, the infimum in (3.3) is attained at $\tilde{\gamma}$. We now use the fact that for a lower semicontinuous, convex, differentiable function f on $\mathbb{R}$,

$$f(x) - xf'(x) = -f^*(f'(x)),$$

where for $v \in \mathbb{R}$ $f^*(v) = \sup_{u \in \mathbb{R}}\{uv - f(u)\}$ is the Legendre-Fenchel transform of f. According to (2.2) $L^i = (H^i)^*$, and so $H^i = (L^i)^*$. We now calculate

$$\begin{aligned}(\lambda^i_\beta)'(\gamma) &= -\tilde{L}^0(0) + L^i(\beta/\gamma) - (\beta/\gamma) \cdot (L^i)'(\beta/\gamma) \\ &= -\tilde{L}^0(0) - (L^i)^*((L^i)'(\beta/\gamma)) \\ &= -\tilde{L}^0(0) - H^i((L^i)'(\beta/\gamma)) .\end{aligned} \tag{3.6}$$

Thus $(\lambda^i_\beta)'(\tilde{\gamma}) = 0$ if and only if $-H^i((L^i)'(b)) = \tilde{L}^0(0)$, where $b = \beta/\tilde{\gamma}$. This completes the proof of the lemma. □

We now present examples of rate functions L in Theorem 2.1 in the case $d = 1$. In each example, there exists a number $c \leqq 0$ such that $\tilde{L}^1(\beta)$ equals $L^1(\beta)$ for $\beta < c$ and (when $c < 0$) $\tilde{L}^1(\beta)$ is an affine function of $\beta \in [c, 0)$. For $i = 1$ the infimum in (3.3) for $\beta < c$ is attained at $\tilde{\gamma} = 1$; for $\beta \in [c, 0)$ the infimum is attained at $\gamma = \beta/c$. Similar remarks apply to the form of $\tilde{L}^2(\beta)$. For $i = 1, 2$ $\bar{\beta}^i$ denotes the mean $\int_{\mathbb{R}} y\mu^i(dy)$, which is the unique minimum point and the unique zero of L^i.

Examples 3.2

(a) *Stable boundary.* We assume that $\bar{\beta}^1$ is positive and $\bar{\beta}^2$ is negative. Thus in each open halfline the random walk tends to move toward the origin. Since for $i = 1, 2$, $L^i(\bar{\beta}^i) = 0$, (2.1) yields $\tilde{L}^0(0) = 0$. For $i = 1, 2$ and $\beta \in \mathbb{R}^d$ satisfying $(-1)^i \beta > 0$, the function $\lambda^i_\beta(\gamma)$ in (3.4) equals β times the slope of the line connecting the points (0,0) and $(\beta/\gamma, L^i(\beta/\gamma))$. There exist unique points $b^1 < 0$ and $b^2 > 0$ such that

$$L^1(b^1)/b^1 = \max_{\beta<0}\{L^1(\beta)/\beta\} \quad \text{and} \quad L^2(b^2)/b^2 = \min_{\beta>0}\{L^2(\beta)/\beta\} .$$

It follows from (3.3) that for $\beta < b^1$ $\tilde{L}^1(\beta)$ equals $L^1(\beta)$; for $\beta \in [b^1, 0)$ $\tilde{L}^1(\beta)$ equals the linear function $\beta L^1(b^1)/b^1$. Similarly, for $\beta > b^2$ $\tilde{L}^2(\beta)$ equals $L^2(\beta)$; for $\beta \in (0, b^2]$ $\tilde{L}^2(\beta)$ equals the linear function $\beta L^2(b^2)/b^2$. The rate function $L(\beta)$ given by (2.4) is shown in Fig. 1. It is convex but not strictly convex, has a unique minimum point at the origin, and is not differentiable at the origin.

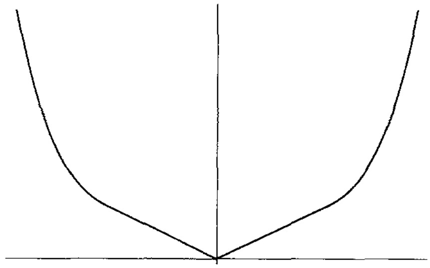

Fig. 1. Rate function for $d = 1$: stable boundary

(b) *Unstable boundary.* We assume that $\bar{\beta}^1$ is negative and $\bar{\beta}^2$ is positive. Thus in each open halfline the random walk tends to move away from the origin. The assumption on $\bar{\beta}^1$ and $\bar{\beta}^2$ implies that

$$\inf_{\beta \geqq 0} L^1(\beta) = L^1(0) > 0 \qquad \text{and} \qquad \inf_{\beta \leqq 0} L^2(\beta) = L^2(0) > 0 .$$

Formula (2.1) yields $\tilde{L}^0(0) = \min[L^1(0), L^2(0)] > 0$.

We first consider the case where $L^1(0) = L^2(0)$. Then $\tilde{L}^0(0) = L^1(0) = L^2(0)$. For $i = 1, 2$ and $\beta \in \mathbb{R}^d$ satisfying $(-1)^i \beta > 0$, the function $\lambda^i_\beta(\gamma)$ in (3.4) equals $L^i(0)$ plus β times the slope of the line connecting the points $(0, L^i(0))$ and $(\beta/\gamma, L^i(\beta/\gamma))$. For β negative (resp., positive), this slope is a monotonically increasing (resp., decreasing) function of $\gamma \in (0, 1]$. It follows from (3.3) that for $\beta < 0$ $\tilde{L}^1(\beta)$ equals $L^1(\beta)$ and for $\beta > 0$ $\tilde{L}^2(\beta)$ equals $L^2(\beta)$. The rate function $L(\beta)$ given by (2.4) is shown in Fig. 2. It is continuous but not convex, has minimum points at $\bar{\beta}^1$ and $\bar{\beta}^2$, and is not differentiable at the origin. $L(\beta)$ is real analytic for $\beta \in (-\infty, 0) \cup (0, \infty)$.

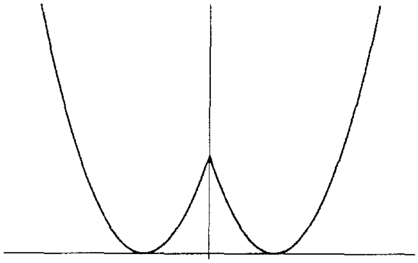

Fig. 2. Rate function for $d = 1$: unstable boundary with $L^1(0) = L^2(0)$

We now consider the case where $L^2(0) > L^1(0)$. By (2.1) $\tilde{L}^0(0) = L^1(0)$. For $\beta < 0$ the function $\lambda^1_\beta(\gamma)$ equals $L^1(0)$ plus β times the slope of the line connecting the points $(0, L^1(0))$ and $(\beta/\gamma, L^1(\beta/\gamma))$. Since the slope is a monotonically increasing function of $\gamma \in (0, 1]$, it follows from (3.3) that for $\beta < 0$ $\tilde{L}^1(\beta)$ equals $L^1(\beta)$. In order to determine the form of $\tilde{L}^2(\beta)$ for $\beta > 0$, we use part (b) of Lemma 3.1. We claim that there exists a unique point $b > 0$ in $(0, \bar{\beta}^2)$ such that

$$-H^2((L^2)'(b)) = \tilde{L}^0(0) = L^1(0) \,. \tag{3.7}$$

Indeed, for $\beta > \bar{\beta}^2$ $(L^2)'(\beta) > 0$. Since $H^2(\alpha) > 0$ for $\alpha > 0$, (3.7) cannot be satisfied for any $b > \bar{\beta}^2$. On the other hand, since $H^2(\alpha)$ attains its minimum value of $-L^2(0)$ at the unique point $\alpha = (L^2)'(0) < 0$, $H^2(\alpha)$ is a monotonically increasing function for $\alpha \in ((L^2)'(0), 0)$; its range is the interval $(-L^2(0), 0)$. Since $L^2(0) > L^1(0) > 0$, we conclude that (3.7) has a unique solution lying in the interval $(0, \bar{\beta}^2)$. For $\beta > b$ $\tilde{L}^2(\beta)$ equals $L^2(\beta)$; for $\beta \in (0, b]$ $\tilde{L}^2(\beta)$ equals to the affine function $L^1(0) + \beta[L^2(b) - L^1(0)]/b$. The rate function $L(\beta)$ given by (2.4) is shown in Fig. 3. It is continuous but not convex, has minimum points at $\bar{\beta}^1$ and $\bar{\beta}^2$, and is not differentiable at the origin. $L(\beta)$ is real analytic for $\beta \in (-\infty, 0)$ and is $\mathscr{C}^1$ but not $\mathscr{C}^2$ for $\beta \in (0, \infty)$.

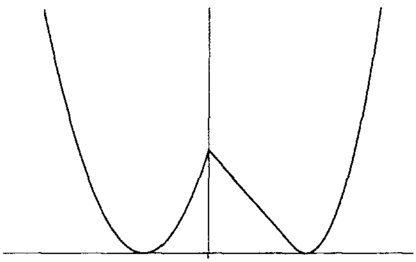

Fig. 3. Rate function for $d = 1$: unstable boundary with $L^2(0) > L^1(0)$

(c) *One-sided unstable boundary.* We assume that $\overline{\beta}^1$ equals 0 and $\overline{\beta}^2$ is positive. Thus $L^1(0) = 0 < L^2(0)$ and

$$\tilde{L}^0(0) = \min\{L^1(0), L^2(0)\} = L^1(0) = 0 .$$

As in the previous example, it follows from (3.3) that for $\beta < 0$ $\tilde{L}^1(\beta)$ equals $L^1(\beta)$. In order to determine the form of $\tilde{L}^2(\beta)$ for $\beta > 0$, we use part (b) of Lemma 3.1. Since $b = \overline{\beta}^2$ is the unique (positive) point satisfying

$$-H^2((L^2)'(b)) = \tilde{L}^0(0) = 0 ,$$

it follows that for $\beta > \overline{\beta}^2$ $\tilde{L}^2(\beta)$ equals $L^2(\beta)$ and for $\beta \in (0, \overline{\beta}^2]$ $\tilde{L}^2(\beta)$ equals 0. The rate function $L(\beta)$ given by (2.4) is shown in Fig. 4. It is differentiable and convex but not strictly convex, and it attains its minimum at all points in the interval $[0, \overline{\beta}^2]$.

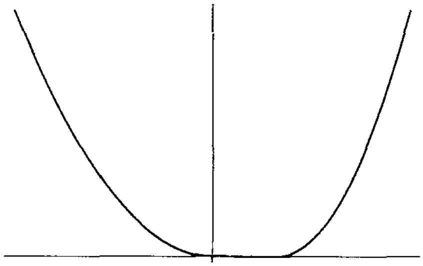

Fig. 4. Rate function for $d = 1$: one-sided unstable boundary

This completes our discussion of Theorem 2.1 and our presentation of examples. The next section presents facts in convex analysis needed in the proof of Theorem 2.1.

4 Convex analysis

The main result in this section is Proposition 4.3, which proves facts about the function L appearing in Theorem 2.1. We follow Rockafellar [28] except that we use the term "convex function" where he uses the term "proper convex function." The Legendre-Fenchel transform f^* of a lower semicontinuous, convex function f on $\mathbb{R}^d$ is defined by the formula

$$f^*(y) = \sup_{x \in \mathbb{R}^d} \{\langle x, y \rangle - f(x)\} \qquad \text{for } y \in \mathbb{R}^d .$$

Let f be a lower semicontinuous, convex function on $\mathbb{R}^d$. We say that f is *superlinear* if $\inf\{x \in \mathbb{R}^d : \|x\| = c\}\ f(x)/c \to \infty$ as $c \to \infty$. The elementary proof of the following lemma is omitted.

Lemma 4.1 *Let f be a lower semicontinuous, convex function on $\mathbb{R}^d$. Then f is superlinear if and only if $f^*(y)$ is finite for every $y \in \mathbb{R}^d$.*

In part (a) of the next proposition, for A a subset of $\mathbb{R}^d$, $\delta(y|A)$ denotes the function that is 0 for $y \in A$ and is $+\infty$ for $y \in \mathbb{R}^d \setminus A$.

Proposition 4.2 *Let f^1 and f^2 be finite convex functions on $\mathbb{R}^d$ (thus automatically lower semicontinuous). We denote by $\phi^i, i = 1, 2$, the respective Legendre-Fenchel transforms of $f^i, i = 1, 2$. Let A^1 and A^2 be nonempty closed convex cones in $\mathbb{R}^d$. We define for $y \in \mathbb{R}^d$*

$$\begin{aligned}\phi(y) = \inf\{&\varrho^1\phi^1(y^1) + \varrho^2\phi^2(y^2) : \varrho^1 \geqq 0,\ \varrho^2 \geqq 0, \\ &\varrho^1 + \varrho^2 = 1,\ y^1 \in A^1, y^2 \in A^2, \\ &\varrho^1 y^1 + \varrho^2 y^2 = y\}\,.\end{aligned} \tag{4.1}$$

Then the following conclusions hold.

(a) *For $y \in \mathbb{R}^d$ $\phi(y)$ may be represented as the Legendre-Fenchel transform*

$$\phi(y) = \sup_{x \in \mathbb{R}^d} \{\langle x, y\rangle - \max[\hat{f}^1(x), \hat{f}^2(x)]\}\,, \tag{4.2}$$

where for $i = 1, 2$ and $x \in \mathbb{R}^d$ $\hat{f}^i(x) = (\phi^i(\cdot) + \delta(\cdot|A^i))^*(x)$.
(b) *The function ϕ is lower semicontinuous, convex, and superlinear, and its Legendre-Fenchel transform ϕ^* is lower semicontinuous, convex, and finite on $\mathbb{R}^d$.*

Proof. (a) For $i = 1, 2$ we define $\hat{\phi}^i(\cdot)$ to be the lower semicontinuous, convex function $\phi^i(\cdot) + \delta(\cdot|A^i)$ and $\hat{f}^i$ to be its Legendre-Fenchel transform. Note that $\hat{\phi}^i \geqq \phi^i$. Then for $y \in \mathbb{R}^d$

$$\begin{aligned}\phi(y) = \inf\{&\varrho^1\hat{\phi}^1(y^1) + \varrho^2\hat{\phi}^2(y^2) : \varrho^1 \geqq 0,\ \varrho^2 \geqq 0, \\ &\varrho^1 + \varrho^2 = 1,\ \varrho^1 y^1 + \varrho^2 y^2 = y\}\,.\end{aligned} \tag{4.3}$$

Since $\hat{f}^i = (\hat{\phi}^i)^* \leqq (\phi^i)^* = f^i$, which is finite on $\mathbb{R}^d$, Theorem 16.5 in Rockafellar [28] shows that

$$\phi(y) = \sup_{x \in \mathbb{R}^d} \{\langle x, y\rangle - \max[\hat{f}^1(x), \hat{f}^2(x)]\}\,. \tag{4.4}$$

This proves part (a).

(b) For $i = 1, 2$ the assumptions on f^i imply that the function $x \mapsto \max[\hat{f}^1(x), \hat{f}^2(x)]$ is lower semicontinuous, convex, and finite on $\mathbb{R}^d$. Since ϕ is the Legendre-Fenchel transform of this function, which in turn is the Legendre-Fenchel of ϕ [see (4.2)], part (b) is a consequence of Lemma 4.1. This completes the proof of Proposition 4.2. □

We now use part (b) of Proposition 4.2 in order to prove that the function $L(\beta)$ appearing in Theorem 2.1 has compact level sets. We also obtain the other properties of this function that are mentioned in part (b) of Theorem 2.1.

Proposition 4.3 *For $\beta \in \mathbb{R}^d$ the function $L(\beta)$ defined in* (2.1)–(2.4) *is lower semicontinuous and superlinear and therefore has compact level sets. In addition $L(\beta)$ is convex on each of the halfspaces $\{\beta \in \mathbb{R}^d : (-1)^i \beta_1 \geqq 0\}$, $i = 1, 2$.*

Proof. The function $\tilde{L}^0(\beta)$ has been defined in (2.1) for $\beta \in \mathbb{R}^d$ satisfying $\beta_1 = 0$. However, the right side of (2.1) is well-defined for all $\beta \in \mathbb{R}^d$. We used the same notation $\tilde{L}^0(\beta)$ to denote the corresponding function on $\mathbb{R}^d$. For $i = 1, 2$ the function $\tilde{L}^i(\beta)$ has been defined in (2.3) for all $\beta \in \mathbb{R}^d$.

The function $\tilde{L}^0$ may be represented as in (4.1) with $\phi^i = L^i = (H^i)^*$ for $i = 1, 2$. Since each H^i is a finite convex function on $\mathbb{R}^d$, part (b) of Proposition 4.2 implies that $\tilde{L}^0$ is lower semicontinuous, convex, and superlinear and that its Legendre-Fenchel transform $(\tilde{L}^0)^*$ is lower semicontinuous, convex, and finite on $\mathbb{R}^d$.

We claim that for $i = 1, 2$ and $\beta \in \mathbb{R}^d$ satisfying $\beta_1 = 0$, $\tilde{L}^i(\beta)$ equals $\tilde{L}^0(\beta)$. Indeed, for such β we have from the definition (2.3) that $\tilde{L}^0(\beta) \geqq \tilde{L}^i(\beta)$. On the other hand, for such β we have from the definition (2.1) that $\tilde{L}^0(\beta) \leqq L^i(\beta)$. Since in the constraints appearing in (2.3) $\gamma^i(\beta^i)_1 = \beta_1 - \gamma^0(\beta^0)_1 = 0$, we obtain a lower bound on $\tilde{L}^i(\beta)$ by replacing $L^i(\beta^i)$ in (2.3) by $\tilde{L}^0(\beta^i)$. The convexity $\tilde{L}^0$ implies that

$$\gamma^0 \tilde{L}^0(\beta^0) + \gamma^i \tilde{L}^0(\beta^i) \geqq \tilde{L}^0(\beta) .$$

We conclude that $\tilde{L}^0(\beta) \geqq \tilde{L}^i(\beta) \geqq \tilde{L}^0(\beta)$ and so the functions must be equal.

It follows that the definition (2.4) of $L(\beta)$ may be replaced by the equivalent definition

$$L(\beta) = \begin{cases} \tilde{L}^1(\beta) & \text{if } \beta_1 \leqq 0 \\ \tilde{L}^2(\beta) & \text{if } \beta_1 \geqq 0 . \end{cases} \tag{4.5}$$

For $i = 1, 2$ the function $\tilde{L}^i$ may be represented as in (4.1) with $\phi^i = L^i = (H^i)^*$ and $\phi^{3-i} = \tilde{L}^0 = ((\tilde{L}^0)^*)^*$. Since H^i and $(\tilde{L}^0)^*$ are finite convex functions on $\mathbb{R}^d$, part (b) of Proposition 4.2 implies that $\tilde{L}^i$ is lower semicontinuous, convex, and superlinear. It now follows from (4.5) that L is lower semicontinuous and superlinear. Since a lower semicontinuous, superlinear function has compact level sets, the first assertion in the proposition is proved. That $L(\beta)$ is convex on each of the halfspaces $\{\beta \in \mathbb{R}^d : (-1)^i \beta_1 \geqq 0\}$, $i = 1, 2$, also follows from (4.5). □

We end this subsection by presenting a consequence of Hypothesis (H) concerning the sets $\text{ri}(\text{dom}\, L^1)$ and $\text{ri}(\text{dom}\, L^2)$, the relative interiors of the effective domains of the functions L^1 and L^2. This lemma will be used several times in the sequel.

Lemma 4.4 *Let μ^1 and μ^2 be Borel probability measures on $\mathbb{R}^d$ that have finite cumulant generating functions and that satisfy Hypothesis (H). Then for $i = 1, 2$*

$$\text{ri}(\text{dom}\, L^i) \cap \text{int}\, \Lambda^{3-i} \neq \emptyset .$$

Proof. Hypothesis (H) states that for $i = 1, 2$

$$\operatorname{supp} \mu^i \cap \operatorname{int} \Lambda^{3-i} \neq \emptyset .$$

According to Theorems VIII.4.3 – VIII.4.4 in [15] (see also [1]), for $i = 1, 2$ the relative interior of the effective domain of the function L^i equals the relative interior of the convex hull of the support of the measure μ^i; in symbols,

$$\operatorname{ri}(\operatorname{dom} L^i) = \operatorname{ri}(\operatorname{conv}(\operatorname{supp} \mu^i)) .$$

Hence by Hypothesis (H) we have for $i = 1, 2$

$$\operatorname{ri}(\operatorname{dom} L^i) \cap \operatorname{int} \Lambda^{3-i} = \operatorname{ri}(\operatorname{conv}(\operatorname{supp} \mu^i)) \cap \operatorname{int} \Lambda^{3-i} \neq \emptyset ,$$

as claimed. □

This completes the convex analysis section of the paper. In the next section, we prove the large deviation bounds in Theorem 2.1.

5 Proofs of the large deviation bounds

In this section we prove the large deviation bounds in Theorem 2.1. The proofs are divided into the following five parts. For $\beta \in \mathbb{R}^d$ and $\varepsilon > 0$, $B(\beta, \varepsilon)$ denotes the open ball $\{y \in \mathbb{R}^d : \|y - \beta\| < \varepsilon\}$.

(a) *Upper bound for* $\beta_1 = 0$. For all $\beta \in \mathbb{R}^d$ satisfying $\beta_1 = 0$,

$$\limsup_{\varepsilon \to 0} \limsup_{n \to \infty} n^{-1} \log P\{S_n/n \in B(\beta, \varepsilon)\} \leqq -\tilde{L}^0(\beta) .$$

(b) *Lower bound for* $\beta_1 = 0$. For all $\beta \in \mathbb{R}^d$ satisfying $\beta_1 = 0$,

$$\liminf_{\varepsilon \to 0} \liminf_{n \to \infty} n^{-1} \log P\{S_n/n \in B(\beta, \varepsilon)\} \geqq -\tilde{L}^0(\beta) .$$

(c) *Upper bound for* $\beta_1 \neq 0$. For $i = 1, 2$ and all $\beta \in \mathbb{R}^d$ satisfying $(-1)^i \beta_1 > 0$,

$$\limsup_{\varepsilon \to 0} \limsup_{n \to \infty} n^{-1} \log P\{S_n/n \in B(\beta, \varepsilon)\} \leqq -\tilde{L}^i(\beta) .$$

(d) *Lower bound for* $\beta_1 \neq 0$. For $i = 1, 2$ and all $\beta \in \mathbb{R}^d$ satisfying $(-1)^i \beta_1 > 0$,

$$\liminf_{\varepsilon \to 0} \liminf_{n \to \infty} n^{-1} \log P\{S_n/n \in B(\beta, \varepsilon)\} \geqq -\tilde{L}^i(\beta) .$$

(e) *Exponential tightness.* For each $M < \infty$ there exists a compact set K in $\mathbb{R}^d$ such that

$$\limsup_{n \to \infty} n^{-1} \log P\{S_n/n \notin K\} \leqq -M .$$

Parts (a) – (d) are proved in Subsections 5a – 5d, respectively. The exponential tightness is an immediate consequence of part (a) of Lemma 5.1. The upper bounds (a) and (c) together with the exponential tightness in (e) yield the upper large deviation bound

$$\limsup_{n\to\infty} n^{-1}\log P\{S_n/n \in F\} \leqq -\inf_{\beta\in F} L(\beta)$$

for each closed set F in $\mathbb{R}^d$ [7]. The lower bounds (b) and (d) yield the lower large deviation bound

$$\liminf_{n\to\infty} n^{-1}\log P\{S_n/n \in G\} \geqq -\inf_{\beta\in G} L(\beta)$$

for each open set G in $\mathbb{R}^d$.

5a Upper bound for $\beta_1 = 0$

Fix $\beta \in \mathbb{R}^d$ satisfying $\beta_1 = 0$ and $\tilde{L}^0(\beta) < +\infty$. In this subsection, we prove that

$$\limsup_{\varepsilon\to 0}\limsup_{n\to\infty} n^{-1}\log P\{S_n/n \in B(\beta,\varepsilon)\} \leqq -\tilde{L}^0(\beta)\,. \tag{5.1}$$

We omit the routine modifications needed to handle the case when $\tilde{L}^0(\beta) = +\infty$.

For $n \in \mathbb{N}$ and $i = 1, 2$, we define the normalized sums

$$r_n^i = \frac{1}{n}\sum_{j=1}^n 1_{\{S_{j-1}\in\Lambda^i\}} \quad\text{and}\quad v_n^i = \frac{1}{n}\sum_{j=1}^n 1_{\{S_{j-1}\in\Lambda^i\}} X_j^i\,. \tag{5.2}$$

The quantity r_n^i represents the fraction of time between 0 and $n-1$ that the random walk spends in the halfspace Λ^i. Note that $r_n^1 + r_n^2 = 1$ and $S_n/n = v_n^1 + v_n^2$. Fix $\gamma \in (0, 1)$ and let $\tau \in (0, 1/2)$ be a number to be specified below. For any $\varepsilon > 0$, we write

$$\{S_n/n \in B(\beta,\varepsilon)\} = A_n^1(\varepsilon,\tau) \cup A_n^2(\varepsilon,\tau) \cup A_n^3(\varepsilon,\tau)\,,$$

where

$$A_n^i(\varepsilon,\tau) = \{S_n/n \in B(\beta,\varepsilon),\ r_n^i \leqq \tau\},\quad i = 1, 2\,,$$

and

$$A_n^3(\varepsilon,\tau) = \{S_n/n \in B(\beta,\varepsilon), r_n^1 > \tau, r_n^2 > \tau\}\,.$$

We prove (5.1) by showing that for all sufficiently small $\varepsilon > 0$ and sufficiently small $\tau \in (0, 1/2)$

$$\limsup_{n\to\infty} n^{-1}\log P\{A_n^i(\varepsilon,\tau)\} \leqq -\tilde{L}^0(\beta) + \gamma,\quad i = 1, 2\,, \tag{5.3}$$

and

(5.4) $$\limsup_{n\to\infty} n^{-1}\log P\{A_n^3(\varepsilon,\tau)\} \leqq -\tilde{L}^0(\beta)+\gamma .$$

We prove (5.4), then (5.3).

The first step is to obtain an exponential bound on v_n^i. This is given in part (a) of the next lemma. This yields the exponential tightness of $\{S_n/n\}$, stated in part (b).

Lemma 5.1 (a) *For any number $M > 0$, there exist positive numbers C and λ such that for $i = 1, 2$ and all $n \in \mathbb{N}$*

$$P\{\|v_n^i\| > \lambda\} \leqq Ce^{-Mn} .$$

(b) *The random vectors $\{S_n/n, n \in \mathbb{N}\}$ are exponentially tight.*

Proof. (a) For $i = 1, 2$ and $n \in \mathbb{N}$, $\|v_n^i\| \leqq n^{-1}\sum_{j=1}^n \|X_j^i\|$. Since for all numbers $\alpha > 0$ $E\{\exp(a\|X_j^i\|)\} < \infty$, part (a) follows from Cramér's Theorem.

(b) This is an immediate consequence of part (a). □

The second step is to deduce a restriction on the set of vectors that the random walk tracks in the two halfspaces. We call the next lemma the "tracking" lemma.

Lemma 5.2 *For any $n \in \mathbb{N}$ and any $\varepsilon > 0$, if $S_n/n \in B(\beta, \varepsilon)$, then $v_n^1 \geqq -\varepsilon$ and $v_n^2 \leqq \varepsilon$.*

Proof. We use the function $f(x) = |x_1|$ as a Lyapunov function. If $S_n/n \in B(\beta, \varepsilon)$, then by considering separately the cases $(S_j)_1 \geqq 0$ and $(S_j)_1 < 0$, we obtain

$$\begin{aligned}
\varepsilon \geqq |(S_n)_1/n| &= \sum_{j=0}^{n-1}(|(S_{j+1})_1/n| - |(S_j)_1/n|) \\
&= \frac{1}{n}\sum_{j=0}^{n-1}(-1_{\{S_j\in\Lambda^1\}}(X_j^1)_1 + 1_{\{S_j\in\Lambda^2\}}(X_j^2)_1) \\
&\quad + \frac{2}{n}\sum_{j=0}^{n-1}(1_{\{S_j\in\Lambda^1\}}[(S_j+X_j^1)_1 \vee 0] - 1_{\{S_j\in\Lambda^2\}}[(S_j+X_j^2)_1 \wedge 0]) \\
&\geqq -(v_n^1)_1 + (v_n^2)_1 .
\end{aligned}$$

By assumption, $|(v_n^1)_1 + (v_n^2)_1| \leqq \varepsilon$. Therefore

$$2(v_n^1)_1 = ((v_n^1)_1 - (v_n^2)_1) + ((v_n^1)_1 + (v_n^2)_1) \geqq -2\varepsilon$$

and

$$2(v_n^2)_1 = ((v_n^2)_1 - (v_n^1)_1) + ((v_n^2)_1 + (v_n^1)_1) \leqq 2\varepsilon .$$

This completes the proof. □

We now proceed with the proof of (5.4). For any numbers $\varepsilon > 0$ and $\lambda > 0$, the tracking lemma, Lemma 5.2, implies that there exists a finite collection of pairs of vectors $\{(\beta_j^1, \beta_j^2) \in \mathbb{R}^d \times \mathbb{R}^d, j = 1, 2, \ldots, \Gamma\}$ satisfying $\beta_j^1 + \beta_j^2 = \beta$, $(\beta_j^1)_1 \geqq 0$, $(\beta_j^2)_1 \leqq 0$, and

$$\left\{\frac{S_n}{n} \in B(\beta, \varepsilon), \|v_n^i\| \leqq \lambda \quad \text{for } i = 1, 2\right\} \subset \bigcup_{j=1}^{\Gamma} \{v_n^i \in B(\beta_j^i, 2\varepsilon) \quad \text{for } i = 1, 2\} .$$

Choosing $N = N(\tau)$ to be the integer that satisfies $\tau^{-1} - 2 \leqq N < \tau^{-1} - 1$, we have

$$\{r_n^1 > \tau, r_n^2 > \tau\} = \{\tau < r_n^1 \leqq 1 - \tau\} \subset \bigcup_{k=1}^{N} \{r_n^1 \in [k\tau, (k+1)\tau]\} \subset (0, 1) .$$

Given $M > 0$ and picking $C > 0$ and $\lambda > 0$ in accordance with Lemma 5.1(a), we have

$$\begin{aligned} (5.5) \quad & P\{A_n^3(\varepsilon, \tau)\} \\ & \leqq P\left\{\frac{S_n}{n} \in B(\beta, \varepsilon), \|v_n^i\| \leqq \lambda \quad \text{for } i = 1, 2 ; \tau < r_n^1 \leqq 1 - \tau\right\} \\ & \quad + \sum_{i=1}^{2} P\{\|v_n^i\| > \lambda\} \\ & \leqq \sum_{j=1}^{\Gamma} \sum_{k=1}^{N} P\{v_n^i \in B(\beta_j^i, 2\varepsilon) \quad \text{for } i = 1, 2 ; r_n^1 \in [k\tau, (k+1)\tau]\} \\ & \quad + 2Ce^{-Mn} . \end{aligned}$$

In order to estimate the probabilities on the last line of (5.5), we introduce a sequence of changes of measure. Let $(\Omega, \mathscr{F}, P)$ denote the probability space on which the random variables $\{X_j^i, j \in \mathbb{N}, i = 1, 2\}$ are defined. Denote by P_n the probability measure on $(\Omega, \mathscr{F})$ induced by the marginal distribution of $\{X_j^i, j = 1, \ldots, n, i = 1, 2\}$ with respect to P. Given $\alpha^1 \in \mathbb{R}^d$, $\alpha^2 \in \mathbb{R}^d$, and $n \in \mathbb{N}$, we define the new measure $P_n^{\alpha^1, \alpha^2}(d\omega)$ on $(\Omega, \mathscr{F})$ with the Radon-Nikodym derivative

$$(5.6) \qquad \frac{dP_n^{\alpha^1, \alpha^2}}{dP_n}(\omega) = \exp\left(n \sum_{i=1}^{2} \langle \alpha^i, v_n^i(\omega)\rangle - n \sum_{i=1}^{2} r_n^i(\omega) H^i(\alpha^i)\right) .$$

A straightforward argument involving conditioning shows that $P_n^{\alpha^1, \alpha^2}$ is a probability measure.

For $i = 1, 2$, $R > 0$, $\gamma \in \mathbb{R}^d$, $\beta^1 \in \mathbb{R}^d$, and $\beta^2 \in \mathbb{R}^d$, we define the functions

$$L_R^i(\gamma) = \sup_{\{\alpha \in \mathbb{R}^d : \|\alpha\| \leqq R\}} \{\langle \alpha, \gamma \rangle - H^i(\alpha)\},$$

$$L_R(\beta^1, \beta^2) = \inf_{0<t<1} \left\{ t L_R^1\left(\frac{\beta^1}{t}\right) + (1-t) L_R^2\left(\frac{\beta^2}{1-t}\right) \right\},$$

and

$$L_R(\beta) = \inf\{L_R(\beta^1, \beta^2) : \beta^1 \in \mathbb{R}^d, \ \beta^2 \in \mathbb{R}^d, \\ \beta^1 + \beta^2 = \beta, \ (\beta^1)_1 \geqq 0, \ (\beta^2)_1 \leqq 0\}.$$

The following lemma relating $L_R(\beta)$ and $\tilde{L}^0(\beta)$ will be proved later in this section.

Lemma 5.3 $\limsup_{R \to \infty} L_R(\beta) \geqq \tilde{L}^0(\beta)$.

The next lemma shows how to bound each of the probabilities on the last line of (5.5).

Lemma 5.4 *Let β^1 and β^2 be vectors in $\mathbb{R}^d$ satisfying $\beta^1 + \beta^2 = \beta$, $(\beta^1)_1 \geqq 0$, $(\beta^2)_1 \leqq 0$, and define for any $R > 0$ the finite number*

$$c(R) = \sup\{|H^1(\alpha)| : \alpha \in \mathbb{R}^d, \ \|\alpha\| \leqq R\}.$$

Then there exists $R > 0$ such that for any $k \in \{1, 2, \ldots, N\}$

$$P\{v_n^i \in B(\beta^i, 2\varepsilon) \quad \text{for } i = 1, 2, \ ; \ r_n^1 \in [k\tau, (k+1)\tau]\} \\ \leqq \exp(-n\tilde{L}^0(\beta) + n\gamma/2 + 4n\varepsilon R + n\tau c(R)).$$

Proof. We prove the lemma under the assumption that $L_R(\beta) < \infty$ for all $R > 0$, omitting the routine modifications needed if this assumption is not true. Given $k \in \{1, 2 \ldots, N\}$, we set $\kappa^1 = k\tau$ and $\kappa^2 = 1 - k\tau$. For any $R > 0$ and any vectors α^1 and α^2 satisfying $\|\alpha^1\| \leqq R$ and $\|\alpha^2\| \leqq R$,

$$P\{v_n^i \in B(\beta^i, 2\varepsilon) \quad \text{for } 1, 2; \ r_n^1 \in [k\tau, (k+1)\tau]\}$$
$$= E_n^{P^{\alpha^1, \alpha^2}} \left\{ \exp\left(-n \sum_{i=1}^{2} [\langle \alpha^i, v_n^i \rangle - r_n^i H^i(\alpha^i)]\right) 1_{\{v_n^i \in B(\beta^i, 2\varepsilon), \, i=1,2\}} 1_{\{r_n^1 \in [k\tau, (k+1)\tau]\}} \right\}$$
$$\leqq \exp(-n[\langle \alpha^1, \beta^1 \rangle - (k+1)\tau H^1(\alpha^1)] - n[\langle \alpha^2, \beta^2 \rangle - (1-k\tau)H^2(\alpha^2)] + 4n\varepsilon R)$$
$$\leqq \exp\left(-n \sum_{i=1}^{2} [\langle \alpha^i, \beta^i \rangle - \kappa^i H^i(\alpha^i)] + 4n\varepsilon R + n\tau |H^1(\alpha^1)|\right).$$

According to Lemma 5.3, there exists $R > 0$ such that $L_R(\beta) \geqq \tilde{L}^0(\beta) - \gamma/2$.

With this value of R, it follows that

$$
\begin{aligned}
&P\{v_n^i \in B(\beta^i, 2\varepsilon) \quad \text{for } i=1,2;\ r_n^1 \in [k\tau, (k+1)\tau]\} \\
&\leqq \exp\left(-n \sum_{i=1}^{2} \kappa^i \sup_{\{\alpha^i \in \mathbb{R}^d : \|\alpha^i\| \leqq R\}} [\langle \alpha^i, \beta^i/\kappa^i \rangle - H^i(\alpha^i)] + 4n\varepsilon R + n\tau c(R) \right) \\
&\leqq \exp(-nL_R(\beta^1, \beta^2) + 4n\varepsilon R + n\tau c(R)) \\
&\leqq \exp(-nL_R(\beta) + 4n\varepsilon R + n\tau c(R)) \\
&\leqq \exp(-n\tilde{L}^0(\beta) + n\gamma/2 + 4n\varepsilon R + n\tau c(R)) .
\end{aligned}
$$

This completes the proof. □

We now conclude the proof of (5.4). Choosing $R > 0$ in accordance with Lemma 5.4, we define

$$
(5.7) \qquad \varepsilon_1 = \gamma/16R \quad \text{and} \quad \tau_1 = [\gamma/4c(R)] \wedge 1/2 ;
$$

whenever $\varepsilon \in (0, \varepsilon_1)$ and $\tau \in (0, \tau_1)$, then $4\varepsilon R + \tau c(R) \leqq \gamma/2$. We now use Lemma 5.4 to bound each of the summands in (5.5). For all $j \in \{1, 2, \dots, \Gamma\}$, $k \in \{1, 2, \dots, N\}$, $\varepsilon \in (0, \varepsilon_1)$, and $\tau \in (0, \tau_1)$, we have

$$
P\{v_n^i \in B(\beta_j^i, 2\varepsilon) \quad \text{for } i=1,2;\ r_n^1 \in [k\tau, (k+1)\tau]\} \leqq \exp(-n\tilde{L}^0(\beta) + n\gamma) ,
$$

and so by (5.5),

$$
P\{A_n^3(\varepsilon, \tau)\} \leqq \Gamma N \exp(-n\tilde{L}^0(\beta) + n\gamma) + 2Ce^{-Mn} .
$$

Choosing M to exceed $\tilde{L}^0(\beta)$, we conclude that the upper bound (5.4) holds.

In order to prove Lemma 5.3, the following auxiliary result is useful.

Lemma 5.5 (a) *If $\zeta_n \to \zeta \in \mathbb{R}^d$, then for $i = 1, 2$ $\liminf_{n\to\infty} L_n^i(\zeta_n) \geqq L^i(\zeta)$.*
(b) *For any sequence of positive real numbers $\{\gamma_n, n \in \mathbb{N}\}$ that converge to ∞, for $i = 1, 2$, and for any $\delta > 0$*

$$
\lim_{n\to\infty} \gamma_n^{-1} \inf\{L_n^i(\zeta) : \zeta \in \mathbb{R}^d, \|\zeta\| \geqq \delta\gamma_n\} = \infty .
$$

Proof. (a) For any $\alpha \in \mathbb{R}^d$ satisfying $\|\alpha\| \leqq n$,

$$
L_n^i(\zeta_n) \geqq \langle \alpha, \zeta_n \rangle - H^i(\alpha) .
$$

The right side converges to $\langle \alpha, \zeta \rangle - H^i(\alpha)$ as $n \to \infty$. Therefore, for all $\alpha \in \mathbb{R}^d$,

$$
\liminf_{n\to\infty} L_n^i(\zeta) \geqq \langle \alpha, \zeta \rangle - H^i(\alpha),
$$

from which part (a) follows.

(b) Take $k \in \mathbb{N}$, $n \in \mathbb{N}$ satisfying $n \geqq k$, ζ any vector in $\mathbb{R}^d$ satisfying $\|\zeta\| \geqq \delta\gamma_n$, and $\alpha = k\zeta/\|\zeta\|$. Then

$$
L_n^i(\zeta) \geqq \langle \alpha, \zeta \rangle - H^i(\zeta) \geqq k\|\zeta\| - c_i(k) \geqq k\delta\gamma_n - c_i(k) ,
$$

where $c_i(k) \doteq \sup\{H^i(\alpha) : \alpha \in \mathbb{R}^d, \|\alpha\| \leqq k\}$. Part (b) is an immediate consequence of these inequalities. □

We now turn to the proof of Lemma 5.3.

Proof of Lemma 5.3. Given $\delta > 0$, we choose sequences $\{t_n, n \in \mathbb{N}\} \subset (0,1)$, $\{b_n^1, n \in \mathbb{N}\} \subset \mathbb{R}^d$, and $\{b_n^2, n \in \mathbb{N}\} \subset \mathbb{R}^d$ satisfying for each $n \in \mathbb{N}$ $(b_n^1)_1 \geqq 0$, $(b_n^2)_1 \leqq 0$, $b_n^1 + b_n^2 = \beta$, and

$$\varrho(n) \doteq t_n L_n^1\left(\frac{b_n^1}{t_n}\right) + (1-t_n)L_n^2\left(\frac{b_n^2}{1-t_n}\right) \leqq L_n(\beta) + \delta \, . \tag{5.8}$$

If $L_n(\beta)$ equals ∞ for infinitely many n or if $\bar{\varrho} \doteq \limsup_{n\to\infty} \varrho(n)$ equals ∞, then the lemma is obvious. Let us therefore assume that neither of these two possibilities occurs.

By passing to subsequences, we may assume without loss of generality that the quantities $\{t_n, n \in \mathbb{N}\}$ converge to a limit $t^* \in [0,1]$. If t^* lies in the open interval $(0,1)$, then by part (b) of Lemma 5.5 and the fact that $\bar{\varrho} < \infty$, it follows that the quantities $\{\|b_n^1\|, n \in \mathbb{N}\}$ and $\{\|b_n^2\|, n \in \mathbb{N}\}$ remain bounded. Hence, there exists a sequence of positive integers $\{n'\}$ converging to ∞ such that both subsequences $\{b_{n'}^1\}$ and $\{b_{n'}^2\}$ converge. In this case, the lemma follows from part (a) of Lemma 5.5 and (5.8).

We now consider the case where $t^* = 0$; the case where $t^* = 1$ is handled similarly. Since $\bar{\varrho} < \infty$, part (b) of Lemma 5.5 excludes the possibility that $\limsup_{n\to\infty} \|b_n^1\| > 0$. Therefore, $\lim_{n\to\infty} \|b_n^1\| = 0$. Using again part (b) of Lemma 5.5 and the fact that $\bar{\varrho} < \infty$, we may choose a subsequence $\{b_{n'}^2\}$ converging to β. It follows from part (a) of Lemma 5.5 that

$$\limsup_{n\to\infty} \varrho(n) \geqq \limsup_{n'\to\infty} (1-t_{n'})L_{n'}^2\left(\frac{b_{n'}^2}{1-t_{n'}}\right) \geqq L^2(\beta) \geqq \tilde{L}^0(\beta) \, .$$

The lemma is a consequence of the last display and the inequality in (5.8). This completes the proof of the lemma. □

The upper bound (5.4) has been proved for all $\varepsilon \in (0, \varepsilon_1)$ and all $\tau \in (0, \tau_1)$. In order to complete the proof of the upper large deviation bound (5.1), we must still prove the upper bound (5.3) for all sufficiently small $\varepsilon \in (0, \varepsilon_1)$ and $\tau \in (0, \tau_1)$. Since $r_n^1 + r_n^2 = 1$ and $S_n/n = v_n^1 + v_n^2$, we may write

$$\begin{aligned} (5.9) \quad & P\{A_n^i(\varepsilon,\tau)\} \\ & \doteq P\{S_n/n \in B(\beta,\varepsilon), r_n^i \leqq \tau\} \\ & = P\{S_n/n \in B(\beta,\varepsilon), r_n^i \leqq \tau, \|v_n^i\| < \varepsilon\} \\ & \quad + P\{S_n/n \in B(\beta,\varepsilon), r_n^i \leqq \tau, \|v_n^i\| \geqq \varepsilon\} \\ & \leqq P\{r_n^{3-i} \geqq 1-\tau, v_n^{3-i} \in \overline{B}(\beta, 2\varepsilon)\} + P\{r_n^i \leqq \tau, \|v_n^i\| \geqq \varepsilon\}, \end{aligned}$$

where $\overline{B}$ denotes the closed ball. The two probabilities on the last line of this display will be handled in the next two lemmas.

Lemma 5.6 *For $i = 1, 2$, there exists $\varepsilon_2 \in (0, \varepsilon_1)$ and $\tau_2 \in (0, \tau_1)$ such that whenever $\varepsilon \in (0, \varepsilon_2)$ and $\tau \in (0, \tau_2)$*

$$\limsup_{n \to \infty} n^{-1} \log P\{r_n^{3-i} \geqq 1 - \tau, v_n^{3-i} \in \overline{B}(\beta, 2\varepsilon)\} \leqq -L^{3-i}(\beta) + \gamma$$
$$\leqq -\tilde{L}^0(\beta) + \gamma .$$

Proof. We carry out the proof under the assumption that $L^1(\beta)$ and $L^2(\beta)$ are finite, omitting the routine modifications needed when either of these quantities equals $+\infty$. We also assume that both $L^1(\beta)$ and $L^2(\beta)$ are positive. If either of these quantities equals zero, then the corresponding upper bound is automatic. For $i = 1, 2$ and any subset A of $\mathbb{R}^d$, we write $L^i(A)$ for the quantity $\inf_{\beta \in A} L^i(\beta)$.

Let $i = 2$, so that $3 - i = 1$. As in the proof of the upper bound in Cramér's theorem (see, e.g., Sect. VII.4 of [15]), for any $\varepsilon > 0$, the compactness of the closed ball $\overline{B}(\beta, 2\varepsilon)$ guarantees that there exist finitely many nonzero vectors $\{\alpha_k^1, k = 1, \ldots, p\}$ such that

$$\begin{aligned}
&P\{r_n^1 \geqq 1 - \tau, v_n^1 \in \overline{B}(\beta, 2\varepsilon)\} \\
&\leqq \sum_{k=1}^{p} P\{r_n^1 \geqq 1 - \tau, \langle \alpha_k^1, v_n^1 \rangle \geqq L^1(\overline{B}(\beta, 2\varepsilon)) + H^1(\alpha_k^1) - \gamma/3\} \\
&\leqq \sum_{k=1}^{p} E^{P_n^{\alpha_k^1, 0}} \{ \exp(-n\langle \alpha_k^1, v_k^1 \rangle \\
&\qquad + n r_n^1 H^1(\alpha_k^1)) 1_{\{r_n^1 \geqq 1 - \tau\}} 1_{\{\langle \alpha_k^1, v_n^1 \rangle \geqq L^1(\overline{B}(\beta, 2\varepsilon)) + H^1(\alpha_k^1) - \gamma/3\}} \} \\
&\leqq \sum_{k=1}^{p} \exp(-nL^1(\overline{B}(\beta, 2\varepsilon)) + n\tau |H^1(\alpha_k^1)| + n\gamma/3) .
\end{aligned}$$

The same proof shows that there exist finitely many nonzero vectors $\{\alpha_\ell^2, \ell = 1, \ldots, q\}$ such that

$$\begin{aligned}
&P\{r_n^2 \geqq 1 - \tau, v_n^2 \in \overline{B}(\beta, 2\varepsilon)\} \\
&\qquad \leqq \sum_{\ell=1}^{q} \exp(-nL^2(\overline{B}(\beta, 2\varepsilon)) + n\tau |H^2(\alpha_\ell^2)| + n\gamma/3) .
\end{aligned}$$

Since for $i = 1, 2$ $\lim_{\varepsilon \to 0^+} L^i(\overline{B}(\beta, 2\varepsilon)) = L^i(\beta)$, there exists $\varepsilon_2 \in (0, \varepsilon_1)$ such that whenever $\varepsilon \in (0, \varepsilon_2)$

$$L^i(\overline{B}(\beta, 2\varepsilon)) \geqq L^i(\beta) - \gamma/3 .$$

Given $\varepsilon \in (0, \varepsilon_2)$, we pick $\tau_2 \in (0, \tau_1)$ so that

$$\tau_2 \times \max_{k=1, \ldots, p; \ell=1, \ldots, q} \{|H^1(\alpha_k^1)|, \ |H^2(\alpha_\ell^2)|\} \leqq \gamma/3 .$$

For $\varepsilon \in (0, \varepsilon_2)$ and $\tau \in (0, \tau_2)$, we have for all $n \in \mathbb{N}$

$$P\{r_n^1 \geqq 1-\tau,\, v_n^1 \in \overline{B}(\beta, 2\varepsilon)\} \leqq p \exp(-nL^1(\beta) + n\gamma)$$

and

$$P\{r_n^2 \geqq 1-\tau,\, v_n^2 \in \overline{B}(\beta, 2\varepsilon)\} \leqq q \exp(-nL^2(\beta) + n\gamma)\ .$$

The lemma follows from these two inequalities and the fact that for $i = 1, 2$ $L^i(\beta) \geqq \widetilde{L}^0(\beta)$. □

The last lemma in this subsection shows how to bound the second probability on the last line of (5.9).

Lemma 5.7 *Given $\varepsilon \in (0, \varepsilon_2)$, there exists $\tau_3 \in (0, \tau_2)$ such that whenever $\tau \in (0, \tau_3)$*

$$\limsup_{n\to\infty} n^{-1} \log P\{r_n^i \leqq \tau,\, \|v_n^i\| \geqq \varepsilon\} \leqq -\widetilde{L}^0(\beta), \quad i = 1, 2\ .$$

Proof. We assume that $\widetilde{L}^0(\beta) > 0$. If $\widetilde{L}^0(\beta) = 0$, then the lemma is automatic.

Let $\{u^r, r = 1, 2, \ldots, d\}$ denote the unit coordinate vectors in $\mathbb{R}^d$ and let N be a positive number to be specified below. Whenever $\|v_n^i\| \geqq \varepsilon$, then for some $r \in \{1, 2, \ldots, d\}$ either $\langle u^r, v_n^i\rangle \geqq \varepsilon/d^{1/2}$ or $-\langle u^r, v_n^i\rangle \geqq \varepsilon/d^{1/2}$. Hence, using estimates similar to those used repeatedly throughout this subsection, we have for any $\tau \in (0, \tau_2)$

$$P\{r_n^i \leqq \tau,\, \|v_n^i\| \geqq \varepsilon\} \leqq \sum_{r=1}^{d} \exp(n\tau|H^i(Nu^r)| - nN\varepsilon/d^{1/2})$$
$$+ \sum_{r=1}^{d} \exp(n\tau|H^i(-Nu^r)| - nN\varepsilon/d^{1/2})\ .$$

We pick $N \in (0, \infty)$ so that $N \times \varepsilon/d^{1/2} = 2\widetilde{L}^0(\beta)$. We then pick $\tau_3 \in (0, \tau_2)$ so that

$$\tau_3 \times \max_{i=1,2,\, r=1,\ldots,d} \{|H^i(Nu^r)|,\ |H^i(-Nu^r)|\} \leqq \widetilde{L}^0(\beta)\ .$$

For this choice of N and for $\tau \in (0, \tau_3)$, we have for all $n \in \mathbb{N}$

$$P\{r_n^i \leqq \tau,\, \|v_n^i\| \geqq \varepsilon\} \leqq 2d \exp[-n\widetilde{L}^0(\beta)]\ .$$

The lemma follows from this inequality. □

The proof of the upper large deviation bound (5.1) is now complete. We next turn to the proof of the lower large deviation bound for balls centered at points $\beta \in \mathbb{R}^d$ satisfying $\beta_1 = 0$.

5b Lower bound for $\beta_1 = 0$

Fix $\beta \in \mathbb{R}^d$ satisfying $\beta_1 = 0$. In this subsection, we prove that

$$\liminf_{\varepsilon \to 0} \liminf_{n \to \infty} n^{-1} \log P\{S_n/n \in B(\beta, \varepsilon)\} \geqq -\tilde{L}^0(\beta) . \tag{5.10}$$

We assume that $\tilde{L}^0(\beta) < \infty$ because otherwise (5.10) is automatic. Several lemmas that are needed are proved at the end of the subsection.

Let $\varepsilon > 0$ and $\gamma > 0$ is given. By the definition of $\tilde{L}^0$, Hypothesis (H), and a continuity property of L^i, $i = 1, 2$, there exist [Lemma 5.8(a)] real numbers ϱ^1 and ϱ^2 and vectors β^1, β^2, and $\bar{\beta}$ satisfying

$$\begin{gathered} \varrho^1 > 0, \;\; \varrho^2 > 0, \;\; \varrho^1 + \varrho^2 = 1, \;\; (\beta^1)_1 > 0, \;\; (\beta^2)_1 < 0, \\ \bar{\beta}_1 = 0, \;\; \varrho^1 \beta^1 + \varrho^2 \beta^2 = \bar{\beta}, \;\; \|\beta - \bar{\beta}\| < \varepsilon, \end{gathered} \tag{5.11}$$

and

$$\varrho^1 L^1(\beta^1) + \varrho^2 L^2(\beta^2) \leqq \tilde{L}^0(\beta) + \gamma . \tag{5.12}$$

In order to prove (5.10), it suffices to show that

$$\liminf_{n \to \infty} n^{-1} \log P\{S_n/n \in B(\bar{\beta}, \varepsilon)\} \geqq -\tilde{L}^0(\beta) - \gamma . \tag{5.13}$$

The key to the proof is again the sequence of changes of measure introduced in the last subsection. The differentiability of H^i, $i = 1, 2$, on $\mathbb{R}^d$ and results in convex analysis [Lemma 5.8(b)] guarantee that for $i = 1, 2$ there exist points $\alpha^i \in \mathbb{R}^d$ satisfying

$$\nabla H^i(\alpha^i) = \beta^i \quad \text{and} \quad L^i(\beta^i) = \langle \alpha^i, \beta^i \rangle - H^i(\alpha^i) . \tag{5.14}$$

Let $(\Omega, \mathscr{F}, P)$ denote the probability space on which the random variables $\{X_j^i, j \in \mathbb{N}, i = 1, 2\}$ are defined. Denote by P_n the probability measure on $(\Omega, \mathscr{F})$ induced by the marginal distribution of $\{X_j^i, j = 1, \ldots, n, i = 1, 2\}$ with respect to P. Recall that for each $n \in \mathbb{N}$ $P_n^{\alpha^1, \alpha^2}(d\omega)$ is the probability measure on $(\Omega, \mathscr{F})$ with the Radon-Nikodym derivative

$$\frac{dP_n^{\alpha^1, \alpha^2}}{dP_n}(\omega) = \exp\left(n \sum_{i=1}^{2} \langle \alpha^i, v_n^i(\omega) \rangle - n \sum_{i=1}^{2} r_n^i(\omega) H^i(\alpha^i) \right) . \tag{5.15}$$

The quantities v_n^i and r_n^i are defined in (5.2). Since $S_n/n = v_n^1 + v_n^2$ and $\varrho^1 \beta^1 + \varrho^2 \beta^2 = \bar{\beta}$, there exists $\varepsilon_1 = \varepsilon_1(\varepsilon) > 0$ such that for all $n \in \mathbb{N}$

$$\bigcap_{i=1,2} [\{v_n^i - r_n^i \beta^i \in B(0, \varepsilon_1)\} \cap \{r_n^i \in B(\varrho^i, \varepsilon_1)\}] \subset \{S_n/n \in B(\bar{\beta}, \varepsilon)\} . \tag{5.16}$$

For $j \in \mathbb{N}$, we denote by $\mathscr{F}_j$ the σ-field generated by the random vectors $\{X^i_\ell, \ell = 1, 2, \ldots, j, i = 1, 2\}$. The measures $\{P_n^{\alpha^1 \alpha^2}, n \in \mathbb{N}\}$ have the following properties for $i = 1, 2$:

(5.17) For $j = 1, \ldots, n$, $E^{P_n^{\alpha^1, \alpha^2}}\{1_{\{S_{j-1} \in \Lambda^i\}} X^i_j | \mathscr{F}_{j-1}\} = 1_{\{S_{j-1} \in \Lambda^i\}} \beta^i$.

(5.18) As $n \to \infty$, $v^i_n - r^i_n \beta^i \to 0$ in probability with respect to $P_n^{\alpha^1, \alpha^2}$.

(5.19) As $n \to \infty$, $r^i_n \to \varrho^i$ in probability with respect to $P_n^{\alpha^1, \alpha^2}$.

The equality in (5.17) follows from the choice of α^i and is elementary to prove. The limit in (5.18) follows from (5.17) and Chebyshev's inequality [Lemma 5.9]. The limit in (5.19) is a consequence of the facts that $(\beta^1)_1 > 0$, $(\beta^2)_1 < 0$, $(\beta^1)_1 \varrho^1 + (\beta^2)_1 \varrho^2 = (\bar{\beta})_1 = 0$ [Lemma 5.10].

We now derive the lower bound (5.13). According to (5.15), the measure P_n is absolutely continuous with respect to $P_n^{\alpha^1, \alpha^2}$. For suitably small $\varepsilon_1 = \varepsilon_1(\varepsilon) > 0$, (5.16) implies that

$$
\begin{aligned}
&P\{S_n/n \in B(\bar{\beta}, \varepsilon)\} \\
&\quad = P_n\{S_n/n \in B(\bar{\beta}, \varepsilon)\} \\
&\quad = \int_{\{S_n/n \in B(\bar{\beta}, \varepsilon)\}} \frac{dP_n}{dP_n^{\alpha^1, \alpha^2}}(\omega) P_n^{\alpha^1, \alpha^2}(d\omega) \\
&\quad \geqq \int_{\cap_{i=1,2}[\{v^i_n - r^i_n \beta^i \in B(0, \varepsilon_1)\} \cap \{r^i_n \in B(\varrho^i, \varepsilon_1)\}]} \frac{dP_n}{dP_n^{\alpha^1, \alpha^2}}(\omega) P_n^{\alpha^1, \alpha^2}(d\omega) .
\end{aligned}
$$

Now $n^{-1} \log(dP_n/dP_n^{\alpha^1, \alpha^2})$ is given by $-\sum_{i=1}^2 [\langle \alpha^i, v^i_n \rangle - r^i_n H^i(\alpha^i)]$, which we rewrite as

$$
(5.20) \qquad -\sum_{i=1}^{2} \{\langle \alpha^i, v^i_n - r^i_n \beta^i \rangle + r^i_n [\langle \alpha^i, \beta^i \rangle - H^i(\alpha^i)]\} .
$$

Combining this with (5.12), (5.14), (5.18), (5.19), and the fact that $\varepsilon_1 > 0$ may be chosen arbitrarily small, we obtain

$$
\begin{aligned}
\liminf_{n \to \infty} n^{-1} \log P\{S_n/n \in B(\bar{\beta}, \varepsilon)\} &\geqq -\sum_{i=1}^{2} \varrho^i [\langle \alpha^i, \beta^i \rangle - H^i(\alpha^i)] \\
&= -[\varrho^1 L^1(\beta^1) + \varrho^2 L^2(\beta^2)] \\
&\geqq -\tilde{L}^0(\beta) - \gamma .
\end{aligned}
$$

This gives the lower bound (5.13) and thus (5.10).

We complete the proof in a series of three lemmas. The first lemma concerns the quantities appearing in (5.11), (5.12), and (5.14).

Lemma 5.8 (a) *Let* $\varepsilon > 0$ $\gamma > 0$, *and* $\beta \in \mathbb{R}^d$ *satisfying* $\beta_1 = 0$ *be given. Then under Hypothesis (H), there exist real numbers* ϱ^1 *and* ϱ^2 *and vectors* β^1, β^2 *and* $\overline{\beta}$ *in* $\mathbb{R}^d$ *satisfying*

$$\varrho^1 > 0,\ \ \varrho^2 > 0,\ \ \varrho^1 + \varrho^2 = 1,\ \ (\beta^1)_1 > 0,\ \ (\beta^2)_1 < 0,$$
$$\overline{\beta}_1 = 0,\ \ \varrho^1\beta^1 + \varrho^2\beta^2 = \overline{\beta},\ \ \|\beta - \overline{\beta}\| < \varepsilon\,,$$

and

$$\varrho^1 L^1(\beta^1) + \varrho^2 L^2(\beta^2) \leqq \widetilde{L}^0(\beta) + \gamma\,.$$

(b) *Under Hypothesis (H), there exists for* $i = 1, 2$, *a point* $\alpha^i \in \mathbb{R}^d$ *satisfying*

$$\nabla H^i(\alpha^i) = \beta^i \quad \text{and} \quad L^i(\beta^i) = \langle \alpha^i, \beta^i \rangle - H^i(\alpha^i)\,,$$

where $\beta^i \in \mathbb{R}^d$ *has the properties given in part* (a).

Proof. (a) By the definition of $\widetilde{L}^0$, there exist real numbers $\widetilde{\varrho}^1$ and $\widetilde{\varrho}^2$ and vectors $\widetilde{\beta}^1$ and $\widetilde{\beta}^2$ in $\mathbb{R}^d$ satisfying

$$\widetilde{\varrho}^1 \geqq 0,\ \ \widetilde{\varrho}^2 \geqq 0,\ \ \widetilde{\varrho}^1 + \widetilde{\varrho}^2 = 1,$$
$$(\widetilde{\beta}^1)_1 \geqq 0,\ \ (\widetilde{\beta}^2)_1 \leqq 0,\ \ \widetilde{\varrho}^1\widetilde{\beta}^1 + \widetilde{\varrho}^2\widetilde{\beta}^2 = \beta\,,$$

and

$$L^1(\widetilde{\beta}^1) < \infty, \quad L^2(\widetilde{\beta}^2) < \infty,$$
$$\widetilde{\varrho}^1 L^1(\widetilde{\beta}^1) + \widetilde{\varrho}^2 L^2(\widetilde{\beta}^2) \leqq \widetilde{L}^0(\beta) + \gamma/2\,.$$

Suppose first that $(\widetilde{\beta}^1)_1 > 0$ and $(\widetilde{\beta}^2)_1 < 0$. Then, since the equations

$$\widetilde{\varrho}^1 + \widetilde{\varrho}^2 = 1 \quad \text{and} \quad \widetilde{\varrho}^1(\widetilde{\beta}^1)_1 + \widetilde{\varrho}^2(\widetilde{\beta}^2)_1 = \beta_1 = 0$$

have the unique solution

$$\widetilde{\varrho}^1 = \frac{-(\widetilde{\beta}^2)_1}{(\widetilde{\beta}^1)_1 - (\widetilde{\beta}^2)_1} > 0 \quad \text{and} \quad \widetilde{\varrho}^2 = \frac{(\widetilde{\beta}^1)_1}{(\widetilde{\beta}^1)_1 - (\widetilde{\beta}^2)_1} > 0\,,$$

part (a) of the lemma holds with $\varrho^1 \doteq \widetilde{\varrho}^1$, $\varrho^2 \doteq \widetilde{\varrho}^2$, $\beta^1 \doteq \widetilde{\beta}^1$, $\beta^2 \doteq \widetilde{\beta}^2$ and $\overline{\beta} \doteq \beta$.

The case where $(\widetilde{\beta}^1)_1 = 0$ and/or $(\widetilde{\beta}^2)_1 = 0$ can be divided into subcases: $\widetilde{\varrho}^1 = 0$; $\widetilde{\varrho}^2 = 0$; $\widetilde{\varrho}^1 > 0$ and $\widetilde{\varrho}^2 > 0$. We will only consider the case $\widetilde{\varrho}^1 = 0$ and note that the arguments for the other cases are similar.

We assume that $\widetilde{\varrho}^1 = 0$, so that $\widetilde{\varrho}^2 = 1$. Then $\widetilde{\beta}^2 = \beta$. For $i = 1, 2$, let b^i be a point in the nonempty convex set $\operatorname{ri}(\operatorname{dom} L^i) \cap \operatorname{int} \Lambda^{3-i}$ [Lemma 4.4]. For any number $\theta \in (0, 1)$, define

$$\beta^1 = b^1 \quad \text{and} \quad \beta^2(\theta) = \widetilde{\beta}^2 + \theta(b^2 - \widetilde{\beta}^2)\,.$$

The point $\beta^2(\theta)$ lies in $\mathrm{ri}(\mathrm{dom}\, L^2) \cap \mathrm{int}\, \Lambda^1$, and since L^2 is convex and lower semicontinuous, $\lim_{\theta\to 0^+} L^2(\beta^2(\theta)) = L^2(\tilde\beta^2)$. We define real numbers $\varrho^1(\theta)$ and $\varrho^2(\theta)$ as the unique solutions of the equations

$$\varrho^1(\theta) + \varrho^2(\theta) = 1 \quad \text{and} \quad \varrho^1(\theta)(\beta^1)_1 + \varrho^2(\theta)(\beta^2(\theta))_1 = 0 \; ;$$

i.e.,

$$\varrho^1(\theta) = \frac{-(\beta^2(\theta))_1}{(\beta^1)_1 - (\beta^2(\theta))_1} > 0, \quad \varrho^2(\theta) = \frac{(\beta^1)_1}{(\beta^1)_1 - (\beta^2(\theta))_1} > 0 \, .$$

We also define

$$\bar\beta(\theta) = \varrho^1(\theta)\beta^1 + \varrho^2(\theta)\beta^2(\theta) \, ,$$

which for any $\theta \in (0, 1)$ satisfies

$$(\bar\beta(\theta))_1 = \varrho^1(\theta)(\beta^1)_1 + \varrho^2(\theta)(\beta^2(\theta))_1 = 0 \, .$$

Since $\lim_{\theta\to 0^+} \varrho^1(\theta) = 0 = \tilde\varrho^1$, $\lim_{\theta\to 0^+} \varrho^2(\theta) = 1 = \tilde\varrho^2$, and $\lim_{\theta\to 0^+} \beta^2(\theta) = \tilde\beta^2 = \beta$, there exists $\theta \in (0, 1)$ sufficiently small such that the quantities $\varrho^1 \doteq \varrho^1(\theta)$, $\varrho^2 \doteq \varrho^2(\theta)$, β^1, and $\beta^2 \doteq \beta^2(\theta)$ satisfy the claims of part (a) of the lemma.

(b) For $i = 1, 2$, the point β^i specified in part (a) of the lemma lies in $\mathrm{ri}(\mathrm{dom}\, L^i)$. Since the function H^i is differentiable on all of $\mathbb{R}^d$, there exists a point $\alpha^i \in \mathbb{R}^d$ satisfying $\nabla H^i(\alpha^i) = \beta^i$ [Corollary 26.4.1 in Rockafellar [28]]. That $L^i(\beta^i) = \langle \alpha^i, \beta^i \rangle - H^i(\alpha^i)$ follows from the definition of L^i as a Legendre-Fenchel transform [Theorem 23.5 in Rockafellar [28]]. □

The limit in (5.18) is proved in the next lemma.

Lemma 5.9 (a) *For $n \in \mathbb{N}$ and $i = 1, 2$,*

$$E^{P_n^{\alpha^1, \alpha^2}} \{ [\| v_n^i - r_n^i \beta^i \|]^2 \} \leqq \frac{1}{n} \sum_{k=1}^{d} \frac{\partial^2 H^i}{\partial [\alpha_k]^2}(\alpha^i) \, .$$

(b) *For $i = 1, 2$*

$$v_n^i - r_n^i \beta^i \to 0 \textit{ in probability with respect to } P_n^{\alpha^1, \alpha^2} \text{ as } n \to \infty \, .$$

Proof. To prove part (a), we define the random vectors

$$Y_j^i = 1_{\{S_{j-1} \in \Lambda^i\}} X_j^i - 1_{\{S_{j-1} \in \Lambda^i\}} \beta^i$$

and denote by $(Y_j^i)_k$, $k = 1, 2, \ldots, d$, the k'th component of Y_j^i. Since for

$1 \leqq j \neq \ell \leqq n \; E^{P_n^{\alpha^1,\alpha^2}}\{(Y_j^i)_k (Y_\ell^i)_k\} = 0$, we have the expansion

$$[n]^2 E^{P_n^{\alpha^1,\alpha^2}}\{[\|v_n^i - r_n^i \beta^i\|]^2\} = \sum_{k=1}^d E^{P_n^{\alpha^1,\alpha^2}}\left\{\left[\sum_{j=1}^n (Y_j^i)_k\right]^2\right\}$$
$$= \sum_{k=1}^d \sum_{j=1}^n E^{P_n^{\alpha^1,\alpha^2}}\{[(Y_j^i)_k]^2\}$$
$$\leqq n \sum_{k=1}^d E^{P_n^{\alpha^1,\alpha^2}}\{[(X_1^i)_k - (\beta^i)_k]^2\} .$$

But for $i = 1, 2$ and $k \in \{1, 2, \ldots, d\}$

$$E^{P_n^{\alpha^1,\alpha^2}}\{[(X_1^i)_k - (\beta^i)_k]^2\} = \frac{\partial}{\partial \alpha_k}\left(\int_{\mathbb{R}^d} x_k e^{\langle \alpha, x\rangle} \mu^i(\mathrm{d}x) \cdot \frac{1}{\int_{\mathbb{R}^d} e^{\langle \alpha, x\rangle}\mu^i(\mathrm{d}x)}\right)\Bigg|_{\alpha=\alpha^i}$$
$$= \frac{\partial^2 H^i}{\partial[\alpha_k]^2}(\alpha^i) .$$

Thus,

$$E^{P_n^{\alpha^1,\alpha^2}}\{[\|v_n^i - r_n^i \beta^i\|]^2\} \leqq \frac{1}{n}\sum_{k=1}^d \frac{\partial^2 H^i}{\partial[\alpha_k]^2}(\alpha^i) ,$$

as claimed. Part (b) follows from part (a) via Chebyshev's inequality. □

The limit in (5.19) is proved in the next lemma.

Lemma 5.10 *For* $i = 1, 2$

$$r_n^i \to \varrho^i \text{ in probability with respect to } P_n^{\alpha^1,\alpha^2} \text{ as } n \to \infty .$$

Proof. According to (5.17), the increments of the random walk $\{1_{\{S_{j-1} \in \Lambda^i\}} X_j^i,$ $j \in \mathbb{N},\; i = 1, 2\}$ satisfy

$$E^{P_n^{\alpha^1,\alpha^2}}\{1_{\{S_{j-1} \in \Lambda^1\}}(X_j^1)_1 | \mathscr{F}_{j-1}\} = 1_{\{S_{j-1} \in \Lambda^1\}}(\beta^1)_1 > 0$$

and

$$E^{P_n^{\alpha^1,\alpha^2}}\{1_{\{S_{j-1} \in \Lambda^2\}}(X_j^2)_1 | \mathscr{F}_{j-1}\} = 1_{\{S_{j-1} \in \Lambda^2\}}(\beta^2)_1 < 0 .$$

Thus, the random walk in each halfspace has a tendency to move toward the boundary. Let $V(x_1) = |x_1|$. The last two displays imply that $V(\cdot)$ can be used as a Lyapunov function to prove the stability of the one-dimensional process $\{(S_j)_1, j \in \mathbb{N}\}$. More precisely, there exist numbers $c > 0$ and $\lambda < \infty$ such that

$$E^{P_n^{\alpha^1,\alpha^2}}\{V((S_{j-1} + 1_{\{S_{j-1} \in \Lambda^1\}} X_j^1 + 1_{\{S_{j-1} \in \Lambda^2\}} X_j^2)_1) - V((S_{j-1})_1) | S_{j-1}\}$$
$$\leqq -c$$

whenever $|(S_{j-1})_1| \geqq \lambda$. A standard result from the stability theory of Markov processes (see, e.g., Sect. 8.4 of [24]), implies that the random variables $\{(S_j)_1, j \in \mathbb{N}\}$ are tight with respect to $P_n^{\alpha^1, \alpha^2}$. Define $r_n(\cdot)$ to be the normalized occupation measures corresponding to the processes $\{(S_j)_1, j \in \mathbb{N}\}$: for any Borel set $B \subset \mathbb{R}$,

$$r_n(B) = \frac{1}{n} \sum_{j=1}^{n} 1_{\{(S_{j-1})_1 \in B\}} .$$

Since the tightness of the set of random variables $\{(S_j)_1, j \in \mathbb{N}\}$ implies the tightness of the sequence of random measures $\{r_n(\cdot), n \in \mathbb{N}\}$ (Theorem 1.6.1 in [26]), for any given $\delta > 0$ there exists $\gamma = \gamma(\delta) < \infty$ such that

$$P_n^{\alpha^1, \alpha^2}\{r_n\{y : |y| \geqq \gamma\} \geqq \delta\} \leqq \delta . \tag{5.21}$$

Given $j \in \mathbb{N}$ and f be a bounded continuous function mapping $\mathbb{R}^d$ to $\mathbb{R}$, we define

$$\mathscr{L} f(S_{j-1}) = E^{P_j^{\alpha^1, \alpha^2}}\{f(S_j) - f(S_{j-1}) | S_{j-1}\} .$$

For $\theta > 0$, let $f_\theta(x) = (x_1 \wedge \theta) \vee (-\theta)$. Since $f_\theta(x)$ has a Lipschitz constant that is independent of θ, there exists $B < \infty$ such that

$$|\mathscr{L} f_\theta(x) - (1_{\Lambda^1}(x)(\beta^1)_1 + 1_{\Lambda^2}(x)(\beta^2)_1)| \leqq B \tag{5.22}$$

for all $\theta > 0$ and $x \in \mathbb{R}^d$. By (5.17), given any $\delta > 0$ and $\gamma < \infty$ there exists $\theta = \theta(\delta, \gamma) > 0$ such that

$$|\mathscr{L} f_\theta(x) - (1_{\Lambda^1}(x)(\beta^1)_1 + 1_{\Lambda^2}(x)(\beta^2)_1)| \leqq \delta \tag{5.23}$$

for all $x \in \mathbb{R}^d$ satisfying $|x_1| \leqq \gamma$. Now let $\delta > 0$ be given and choose $\gamma < \infty$ according to (5.21) and $\theta > 0$ according to (5.23). Temporarily fixing γ and θ, we consider the quantity

(5.24)

$$\frac{1}{n}(f_\theta(S_n) - f_\theta(S_0)) - \frac{1}{n} \sum_{j=1}^{n} \mathscr{L} f_\theta(S_{j-1}) = \frac{1}{n} \sum_{j=1}^{n} [f_\theta(S_j) - f_\theta(S_{j-1}) - \mathscr{L} f_\theta(S_{j-1})] .$$

For $j = 1, \ldots, n$, define

$$\triangle_j = f_\theta(S_j) - f_\theta(S_{j-1}) - \mathscr{L} f_\theta(S_{j-1}) .$$

Since $E^{P_n^{\alpha^1, \alpha^2}}\{\triangle_j | S_{j-1}\} = 0$ and

$$|\triangle_j| = |f_\theta(S_j) - E^{P_j^{\alpha^1, \alpha^2}}\{f_\theta(S_j) | S_{j-1}\}| \leqq 2\|f_\theta\|_\infty ,$$

we may bound the second moment of the quantity on the left side of (5.24) as follows:

$$E^{P_n^{\alpha^1,\alpha^2}}\left\{\left[\frac{1}{n}(f_\theta(S_n)-f_\theta(S_0))-\frac{1}{n}\sum_{j=1}^{n}\mathscr{L}f_\theta(S_{j-1})\right]^2\right\}$$

$$=\frac{1}{[n]^2}E^{P_n^{\alpha^1,\alpha^2}}\left\{\left[\sum_{j=1}^{n}\triangle_j\right]^2\right\}$$

$$=\frac{1}{[n]^2}\sum_{j=1}^{n}E^{P_j^{\alpha^1,\alpha^2}}\{[\triangle_j]^2\}$$

$$\leqq\frac{4}{n}[\|f_\theta\|_\infty]^2 .$$

Since $n^{-1}(f_\theta(S_n)-f_\theta(S_0))\to 0$, Chebyshev's inequality gives

(5.25) $$\frac{1}{n}\sum_{j=1}^{n}\mathscr{L}f_\theta(S_{j-1})\to 0 \text{ in probability with respect to } P_n^{\alpha^1,\alpha^2} .$$

We can write

$$r_n^1(\beta^1)_1+r_n^2(\beta^2)_1=\frac{1}{n}\sum_{j=1}^{n}(1_{\Lambda^1}(S_{j-1})(\beta^1)_1+1_{\Lambda^2}(S_{j-1})(\beta^2)_1)$$

$$=\int_{\mathbb{R}}\left(1_{\{x_1\leqq 0\}}(\beta^1)_1+1_{\{x_1>0\}}(\beta^2)_1\right)r_n(\mathrm{d}x_1) .$$

Using (5.21), (5.22), and (5.23), we conclude that

$$P_n^{\alpha^1,\alpha^2}\left\{\left|\int_{\mathbb{R}}(1_{\{x_1\leqq 0\}}(\beta^1)_1+1_{\{x_1>0\}}(\beta^2)_1)r_n(\mathrm{d}x_1)\right|\right.$$

$$\left.\geqq B\delta+\delta+\left|\frac{1}{n}\sum_{j=1}^{n}\mathscr{L}f_\theta(S_{j-1})\right|\right\}\leqq\delta .$$

Using (5.25) and the fact that $\delta>0$ is arbitrary, we have

$$r_n^1(\beta^1)_1+r_n^2(\beta^2)_1\to 0$$

in probability with respect to $P_n^{\alpha^1,\alpha^2}$.

Thus for all $n\in\mathbb{N}$, the quantities r_n^1 and r_n^2 satisfy

$$r_n^1+r_n^2=1 \quad\text{and}\quad r_n^1(\beta^1)_1+r_n^2(\beta^2)_1=\sigma_n ,$$

where σ_n tends to zero in probability with respect to $P_n^{\alpha^1,\alpha^2}$. Since $(\beta^1)_1 > 0$ and $(\beta^2)_1 < 0$, the equations in the last display have the unique solutions

$$r_n^1 = \frac{-(\beta^2)_1 + \sigma_n}{(\beta^1)_1 - (\beta^2)_1} = \varrho^1 + \frac{\sigma_n}{(\beta^1)_1 - (\beta^2)_1},$$
$$r_n^2 = \frac{(\beta^1)_1 - \sigma_n}{(\beta^1)_1 - (\beta^2)_1} = \varrho^2 - \frac{\sigma_n}{(\beta^1)_1 - (\beta^2)_1}.$$

We conclude that for $i = 1, 2$

$$r_n^i \to \varrho^i \text{ in probability with respect to } P_n^{\alpha^1,\alpha^2} \text{ as } n \to \infty .$$

This completes the proof of the lemma. □

The proof of the lower large deviation bound (5.10) is now complete. We next turn to the proof of the upper large deviation bound for balls centered at points $\beta \in \mathbb{R}^d$ satisfying $\beta_1 \neq 0$.

5c *Upper bound for* $\beta_1 \neq 0$

Fix $\beta \in \mathbb{R}^d$ satisfying $\beta_1 \neq 0$ and $L(\beta) < \infty$. In this subsection, we prove that

$$\limsup_{\varepsilon \to 0} \limsup_{n \to \infty} n^{-1} \log P\{S_n/n \in B(\beta, \varepsilon)\} \leqq -L(\beta) . \tag{5.26}$$

We omit the routine modifications needed to handle the case when $L(\beta) = \infty$. By symmetry, it is enough to treat vectors $\beta \in \mathbb{R}^d$ satisfying $\beta_1 > 0$, in which case $L(\beta) = \tilde{L}^2(\beta) < \infty$.

Fix $\gamma > 0$. We will make use of parameters, ε, η, and τ satisfying $\varepsilon \in (0, \beta_1), \eta \in (0, \beta_1 - \varepsilon)$, and $\tau \in (0, 1)$. These parameters will be further adjusted in the course of the proof.

For $\tau \in (0, 1)$ such that τ^{-1} is an integer and for $\eta \in (0, \beta_1 - \varepsilon)$, we consider the following decomposition of the event $\{S_n/n \in B(\beta, \varepsilon)\}$:

$$\{S_n/n \in B(\beta, \varepsilon)\} = (\cup_{i=0}^{\tau^{-1}-1} E_{n,i}) \cup E_n ,$$

where for $i \in \{0, 1, \ldots, \tau^{-1}-1\}$

$$\begin{aligned} E_{n,i} = \{&(S_{[ni\tau]})_1/n \leqq \eta, (S_{[nj\tau]})_1/n \geqq \eta \quad \text{for all } j \in \{i+1, \ldots, \tau^{-1}\}, \\ &(S_k)_1/n > 0 \quad \text{for all } k \in \{[ni\tau], \ldots, n\}, \ S_n/n \in B(\beta, \varepsilon)\} \end{aligned} \tag{5.27}$$

and

$$E_n = \{S_n/n \in B(\beta, \varepsilon)\} \setminus (\cup_{i=0}^{\tau^{-1}-1} E_{n,i}) .$$

In (5.27), the square brackets denote the integer part. This decomposition will allow us to bound the probability $P\{S_n/n \in B(\beta, \varepsilon)\}$ by probabilities involving the finite collection of random vectors $\{S_{[nj\tau]}, j = 0, 1, \ldots, \tau^{-1}\}$.

In particular, we will see that once $\eta \in (0, \beta_1 - \varepsilon)$ has been fixed, it is possible to choose $\tau \in (0, 1)$ in such a way that the event E_n has negligible probability.

Given $n \in \mathbb{N}$ and $\zeta \in (0, \infty)$, we define the set

$$\widehat{E}_n = \{ \sup_{\kappa \in \{0, 1, \ldots, n\}} \|S_\kappa / n\| \geqq \zeta \} .$$

For $n \in \mathbb{N}$ and $i \in \{0, 1, \ldots, \tau^{-1} - 1\}$, we also define the sets

$$\overline{E}_{n,i} = E_{n,i} \cap \widehat{E}_n^c .$$

Since

$$\{S_n/n \in B(\beta, \varepsilon)\} \subset \left(\bigcup_{i=0}^{\tau^{-1}-1} \overline{E}_{n,i} \right) \cup E_n \cup \widehat{E}_n ,$$

it follows that

$$\text{(5.28)} \quad \limsup_{n \to \infty} n^{-1} \log P\{S_n/n \in B(\beta, \varepsilon)\}$$
$$\leqq \max \left\{ \max_{i \in \{0, 1, \ldots, \tau^{-1} - 1\}} \left(\limsup_{n \to \infty} n^{-1} \log P\{\overline{E}_{n,i}\} \right), \right.$$
$$\left. \limsup_{n \to \infty} n^{-1} \log P\{E_n\}, \limsup_{n \to \infty} n^{-1} \log P\{\widehat{E}_n\} \right\} .$$

We now derive an upper bound for each of the limits superior in this display.

The last limit superior in (5.28) is easily bounded using the following process-level exponential tightness result, Lemma 5.11. For later use, it will be stated in a more general form than is needed here. Lemma 5.11 implies that there exists $\zeta \in (0, \infty)$ such that

$$\text{(5.29)} \quad \limsup_{n \to \infty} n^{-1} \log P\{\widehat{E}_n\} \leqq -\widetilde{L}^2(\beta) + \gamma .$$

Given $N \in (0, \infty)$, we define the compact hypercube

$$K(N) = \{x \in \mathbb{R}^d : |x_i| \leqq N, \ \ i = 1, \ldots, d\} .$$

For $x \in \mathbb{R}^d$, P_x denotes probability conditioned on $S_0 = x$.

Lemma 5.11 *For each $M < \infty$, there exists a number $\lambda = \lambda(M) \in (0, \infty)$ such that for all combinations of $\tau \in (0, 1]$ and $N \in (0, \infty)$ satisfying $N/\tau \geqq \lambda$*

$$\text{(5.30)} \quad \limsup_{n \to \infty} n^{-1} \log P_x \left\{ \sup_{j \in \{0, 1, \ldots, [\tau n]\}} (S_j - x)/n \notin K(N) \right\} \leqq -M ,$$

where $[\tau n]$ denotes the integer part of τn. The inequality (5.30) *holds uniformly over $x \in \mathbb{R}^d$.*

The standard proof follows easily from the fact that for any $\alpha \in \mathbb{R}^d$, the random vectors

$$\exp[\langle \alpha, S_k \rangle - (k r_k^1 H^1(\alpha) + k r_k^2 H^2(\alpha))], \quad k = 1, 2, \ldots, n ,$$

form a P-martingale with mean 1. We omit the details. See Lemma 2.5 in [12] for a continuous-time analogue.

We now bound the middle limit superior in (5.28). The inclusion $\eta \in (0, \beta_1 - \varepsilon)$ implies that for $\omega \in \{S_n/n \in B(\beta, \varepsilon)\}$ there exists an $i \in \{0, 1, \ldots, \tau^{-1} - 1\}$ such that

$$(S_{[ni\tau]})_1/n \leqq \eta, \; (S_{[nj\tau]})_1/n \geqq \eta \quad \text{for all } j \in \{i+1, \ldots, \tau^{-1}\}, \; S_n/n \in B(\beta, \varepsilon) .$$

Thus, if $\omega \in E_n$, then for some $k \in \{[ni\tau], \ldots, n\}$ we must have $(S_k)_1/n \leqq 0$. Therefore, it must be true that

$$|((S_{[n\iota\tau]})_1 - (S_\kappa)_1)/n| \geqq \eta$$

for some ι and κ satisfying

$$[n(\iota - 1)\tau \leqq \kappa \leqq [n\iota\tau] .$$

Lemma 5.11 implies we may choose $\tau_0 = \tau_0(\eta) \in (0, 1)$ so that for $\varepsilon \in (0, \beta_1)$ and $\tau \in (0, \tau_0)$

$$\limsup_{n \to \infty} n^{-1} \log P\{E_n\} \leqq -\tilde{L}^2(\beta) + \gamma . \tag{5.31}$$

The proof of the upper bound (5.26) thus reduces to proving that for each $\tau \in (0, \tau_0)$, for each $i \in \{0, 1, \ldots, \tau^{-1} - 1\}$, and for all sufficiently small $\varepsilon \in (0, \beta_1)$

$$\limsup_{n \to \infty} n^{-1} \log P\{\overline{E}_{n,i}\} \leqq -\tilde{L}^2(\beta) + \gamma . \tag{5.32}$$

We need the following lemma.

Lemma 5.12 *For each $\tau \in (0, \tau_0)$, there exist numbers $\varepsilon_0 = \varepsilon_0(|\beta_1|, \gamma) \in (0, |\beta_1|)$ and $c = c(|\beta_1|, \gamma) \in (0, 1/2)$, both independent of τ, such that*

$$\limsup_{n \to \infty} n^{-1} \log P\{E_{n,i}\} \leqq -L(\beta) + \gamma$$

whenever $\varepsilon \in (0, \varepsilon_0)$ and whenever $i\tau \leqq c$ or $i\tau \geqq 1 - c$.

Proof. We will give the proof assuming that $\beta_1 > 0$, $L(\beta) < \infty$, and $L^2(\beta) < \infty$. The arguments are easily modified if $\beta_1 < 0$, $L(\beta) = \infty$, or $L^2(\beta) = \infty$. Since $\beta_1 > 0$, we have $L(\beta) = \tilde{L}^2(\beta)$.

Case 1. $i\tau \leqq c$. Choose $\delta > 0$ so that

$$\inf_{\beta' \in B(\beta, \delta)} L^2(\beta') \geqq L^2(\beta) - \gamma/2 . \tag{5.33}$$

We may then choose $c_1 \in (0, 1/2)$ so that for all $s \in (0, c_1)$

$$B\left(\frac{\beta}{1-s}, \frac{c_1}{1-s}\right) \subset B(\beta, \delta) \quad \text{and} \quad c_1 L^2(\beta) \leqq \gamma/2 . \tag{5.34}$$

Define $\varepsilon_0 = \min\{c_1/2, \beta_1/2\}$. By Lemma 5.11 we may choose $c_2 \in (0, c_1)$ so that

$$\limsup_{n\to\infty} n^{-1} \log P\left\{ \sup_{k\in\{0,1,\dots,[2nc_2]\}} \|S_k/n\| \geqq c_1/2 \right\} \leqq -L(\beta) + \gamma .$$

If $\varepsilon \in (0, \varepsilon_0)$ and $i\tau \leqq c_2$, then the definition of $E_{n,i}$ and the Markov property imply that

$$(5.35) \qquad \limsup_{n\to\infty} n^{-1} \log P\{E_{n,i}\} \leqq \min\{[-L(\beta)+\gamma], \limsup_{n\to\infty} n^{-1} \log q_n\} ,$$

where

$$\begin{aligned} q_n &= P\left\{ \frac{1}{n} \sum_{j=[ni\tau]+1}^{n} X_j^2 \in B(\beta, c_1) \right\} \\ &= P\left\{ \frac{1}{n(1-i\tau)} \sum_{j=[ni\tau]+1}^{n} X_j^2 \in B\left(\frac{\beta}{1-i\tau}, \frac{c_1}{1-i\tau} \right) \right\} . \end{aligned}$$

Cramér's Theorem, (5.33), and (5.34) yield

$$\begin{aligned} \limsup_{n\to\infty} n^{-1} \log q_n &\leqq -(1-i\tau) \times \inf\left\{ L^2(\beta') : \beta' \in B\left(\frac{\beta}{1-i\tau}, \frac{c_1}{1-i\tau} \right) \right\} \\ &\leqq -L^2(\beta) + \gamma \\ &\leqq -L(\beta) + \gamma , \end{aligned}$$

The last display and (5.35) conclude the proof of Case 1.

Case 2. $i\tau \geqq 1-c$. If $\varepsilon \in (0, \beta_1)$, then the definition of the sets $E_{n,i}$ implies that the process $\{S_k/n\}$ moves a distance at least $\beta_1 - \varepsilon > 0$ in $[n(1-i\tau)]$ units of time. By Lemma 5.11, there exists $c_3 \in (0, 1/2)$ such that

$$\limsup_{n\to\infty} n^{-1} \log P\{E_{n,i}\} \leqq -L(\beta) + \gamma$$

whenever $i\tau \geqq 1 - c_3$. In order to conclude the lemma, we take $c = \min\{c_2, c_2\}$. □

Since each set $\overline{E}_{n,i}$ is a subset of $E_{n,i}$, Lemma 5.12 shows that there exist numbers $\varepsilon_0 = \varepsilon_0(\beta_1, \gamma) \in (0, \beta_1)$ and $c = c(\beta_1, \gamma) \in (0, 1/2)$, both independent of τ, such that (5.32) is true whenever $\varepsilon \in (0, \varepsilon_0)$ and whenever $i\tau \leqq c$ or $i\tau \geqq 1-c$. Fix such a $c \in (0, 1/2)$. We are left to prove (5.32) in the cases where i satisfies $c \leqq i\tau \leqq 1-c$.

Our work in Subsection 5a shows that for every $\overline{\beta}$ in the compact set $\{\overline{\beta} \in \mathbb{R}^d : \overline{\beta}_1 = 0, \|\overline{\beta}\| \leqq K/c\}$, there exists $\delta = \delta(\overline{\beta}) > 0$ such that

$$\limsup_{n\to\infty} n^{-1} \log P\{S_n/n \in B(\overline{\beta}, \delta)\} \leqq -\tilde{L}^0(\overline{\beta}) + \gamma .$$

Fix $\lambda > 0$. An open covering argument applied to the compact set $\{\overline{\beta} \in \mathbb{R}^d : \overline{\beta}_1 = 0, \|\overline{\beta}\| \leqq K/c\}$ produces a finite collection of vectors $\{\beta_v, v = 1, \dots, s\}$

and a associated collection of positive scalars $\{\delta_\nu, \nu = 1, \dots, s\}$ such that the following hold: for all $\nu \in \{1, \dots, s\}$ $(\beta_\nu)_1 = 0$ and $0 < \delta_\nu \leqq \lambda$; for all $\nu \in \{1, \dots, s\}$

$$\limsup_{n\to\infty} n^{-1} \log P\{S_n/n \in B(\beta_\nu, \delta_\nu)\} \leqq -\tilde{L}^0(\beta_\nu) + \gamma \; ; \tag{5.36}$$

there exists $\eta \in (0, 1)$ such that

$$\{\tilde{\beta} \in \mathbb{R}^d : |\tilde{\beta}_1| \leqq (2\eta/c), \|\tilde{\beta}\| \leqq K/c\} \subset \bigcup_{\nu=1}^{s} B(\beta_\nu, \delta_\nu) \, . \tag{5.37}$$

With this choice of η, choose $\tau_0 \in (0, 1)$ so that (5.31) holds for all $\tau \in (0, \tau_0)$.

Now fix a value of i satisfying $c \leqq i\tau \leqq 1 - c$. Let $V = \{v_\ell, \ell = 1, \dots, t\}$ be a finite collection of vectors satisfying $(v_\ell)_1 = 0$ for $\ell \in \{1, \dots, t\}$ and having the property that

$$\begin{aligned} &\{\tilde{\beta} \in \mathbb{R}^d : |\tilde{\beta}_1| \leqq \eta, \|\tilde{\beta}\| \leqq K\} \\ &\subset \bigcup_{\ell=1}^{t} B(v_\ell, 2\eta) \cap \{\tilde{\beta} \in \mathbb{R}^d : \|\tilde{\beta}\| \leqq K\} \\ &\subset \{\tilde{\beta} \in \mathbb{R}^d : |\tilde{\beta}_1| \leqq 2\eta, \|\tilde{\beta}\| \leqq K\} \, . \end{aligned} \tag{5.38}$$

For $n \in \mathbb{N}$ and $\ell \in \{1, \dots, t\}$, we define the sets

$$F_{n,i,\ell} = \{S_{[ni\tau]}/n \in B(v_\ell, 2\eta)\}$$

and

$$\begin{aligned} F_{n,i} = \{(S_{[nj\tau]})_1 \geqq \eta \quad &\text{for all } j \in \{i+1, \dots, \tau^{-1}\}, \\ (S_k)_1/n > 0 \quad &\text{for all } k \in \{[ni\tau], \dots, n\}, S_n/n \in B(\beta, \varepsilon)\} \, , \end{aligned}$$

where the square brackets denote the integer part. The first inclusion in (5.38) implies that

$$\bar{E}_{n,i} \subset \bigcup_{\ell=1}^{t} F_{n,i,\ell} \cap F_{n,i} \, .$$

According to (5.37), since $i\tau \geqq c$, a subcollection of the balls $\{B(\beta_\nu, \delta_\nu)\}$ (say $\nu = \nu_1, \dots, \nu_u$) cover $B(v_\ell/i\tau, 2\eta/c)$ for each $\ell \in \{1, \dots, t\}$. Using (5.36), we derive the upper bounds

$$\begin{aligned} &\limsup_{n\to\infty} n^{-1} \log P\{F_{n,i,\ell}\} \\ &\leqq \limsup_{n\to\infty} n^{-1} P\{S_{[ni\tau]}/n \in B(v_\ell, 2\eta)\} \\ &\leqq \limsup_{n\to\infty} n^{-1} \log P\{S_{[ni\tau]}/ni\tau \in B(v_\ell/i\tau, 2\eta/c)\} \\ &\leqq -i\tau \times \left[\min_{r=1,\dots,u} \tilde{L}^0(\beta_{\nu_r}) - \gamma\right] \\ &\leqq -i\tau \times [\inf\{\tilde{L}^0(\beta^0) : \beta^0 \in B(v_\ell/i\tau, (2n/c) + \lambda), (\beta^0)_1 = 0\} - \gamma] \, . \end{aligned} \tag{5.39}$$

In addition, since $\varepsilon \in (0, \beta_1)$ and $i\tau \leqq 1 - c$, we have the upper bounds

$$\text{(5.40)} \quad \limsup_{n \to \infty} n^{-1} \log P\{F_{n,i} | F_{n,i,\ell}\}$$
$$\leqq \limsup_{n \to \infty} n^{-1} \log P\left\{\frac{1}{n} \sum_{j=[ni\tau]+1}^{n} X_j^2 \in B(\beta - v_\ell, \varepsilon + 2\eta)\right\}$$
$$\leqq -(1 - i\tau) \times \inf\left\{L^2(\beta^2) : \beta^2 \in B\left(\frac{\beta - v_\ell}{1 - i\tau}, \frac{\varepsilon + 2\eta}{1 - i\tau}\right)\right\}$$
$$\leqq -(1 - i\tau) \times \inf\left\{L^2(\beta^2) : \beta^2 \in B\left(\frac{\beta - v_\ell}{1 - i\tau}, \frac{\varepsilon + 2\eta}{c}\right)\right\}.$$

The first inequality is a consequence of the Markov property and the fact that for ω in the set $F_{n,i,\ell}$ the random walk remains in the right halfspace, in which it is spatially homogeneous. The second inequality is a consequence of Cramér's Theorem.

For any value of i satisfying $c \leqq i\tau \leqq 1 - c$ and any $\ell \in \{1, \ldots, t\}$, displays (5.39) and (5.40) imply that a upper bound for each term of the form

$$\limsup_{n \to \infty} n^{-1} \log P\{F_{n,i,\ell} \cap F_{n,i}\}$$

is

$$-\Bigg[i\tau \times \left(\inf\left\{\tilde{L}^0(\beta^0) : \beta^0 \in B\left(\frac{v_\ell}{i\tau}, \frac{2\eta}{c} + \lambda\right), \ (\beta^0)_1 = 0\right\}\right)$$
$$+ (1 - i\tau) \times \left(\inf\left\{L^2(\beta^2) : \beta^2 \in B\left(\frac{\beta - v_\ell}{1 - i\tau}, \frac{\varepsilon + 2\eta}{c}\right)\right\}\right)\Bigg] + \gamma .$$

The latter, in turn, is bounded above by

$$-\inf\left\{\tilde{L}^2(\beta') : \beta' \in B\left(\beta, \frac{\varepsilon + 2\eta}{c} + \lambda\right)\right\} + \gamma .$$

The upper bound (5.26) now follows by sending first $\eta \to 0$, then $\varepsilon \to 0$, then $\lambda \to 0$, and finally $\gamma \to 0$, and by making use of the lower semicontinuity of $\tilde{L}^2$.

To summarize, we pick $\beta \in \mathbb{R}^d$ satisfying $\beta_1 > 0$, $\gamma \in (0, 1)$, and $\lambda \in (0, 1)$. We then make the following choices:

(a) $\zeta = \zeta(\beta, \gamma) \in (0, \infty)$ so that (5.29) holds.
(b) $c = c(\beta_1, \gamma) \in (0, 1/2)$ and $\varepsilon = \varepsilon(\beta_1, \gamma) \in (0, \beta_1)$ so that (5.32) holds.
(c) A finite collection of vectors $\{\beta_\nu, \nu = 1, \ldots, s\}$, where $s = s(\lambda, \zeta, c)$, so that (5.36) and (5.37) hold.
(d) $\eta = \eta(\beta_1, \lambda, \zeta, c, \varepsilon) \in (0, \beta_1 - \varepsilon)$ so that (5.37) holds.
(e) $\tau_0 = \tau_0(\eta) \in (0, 1)$ so that (5.31) holds.
(f) A finite collection of vectors $\{v_\ell, \ell = 1, \ldots, t\}$, where $t = t(\eta, \zeta)$, so that (5.38) holds.

The proof of the upper large deviation bound (5.26) is now complete. We next turn to the proof of the corresponding lower bound.

5d Lower bound for $\beta_1 \neq 0$

Fix $\beta \in \mathbb{R}^d$ satisfying $\beta_1 \neq 0$. In this subsection, we prove the lower bound

$$\liminf_{\varepsilon \to 0} \liminf_{n \to \infty} n^{-1} \log P\{S_n/n \in B(\beta, \varepsilon)\} \geqq -L(\beta) , \tag{5.41}$$

We assume that $L(\beta) < \infty$ because otherwise (5.41) is automatic. By symmetry, it is enough to treat vectors $\beta \in \mathbb{R}^d$ satisfying $\beta_1 > 0$, in which case $L(\beta) = \tilde{L}^2(\beta) < \infty$.

Fix $\varepsilon > 0$ and $\gamma > 0$. By the definition of $\tilde{L}^2(\beta)$, there exist real numbers γ^0 and γ^2 and vectors β^0 and β^2 in $\mathbb{R}^d$ satisfying

$$\gamma^0 \geqq 0, \quad \gamma^2 \geqq 0, \quad \gamma^0 + \gamma^2 = 1,$$
$$(\beta^0)_1 = 0, \quad \gamma^0\beta^0 + \gamma^2\beta^2 = \beta,$$

and

$$\gamma^0 \tilde{L}^0(\beta^0) + \gamma^2 L^2(\beta^2) \leqq \tilde{L}^2(\beta) + \gamma/2 . \tag{5.42}$$

Since $\beta_1 > 0$, we must have $(\beta^2)_1 > 0$, $\gamma^0 < 1$, and $\gamma^2 > 0$. The last inequality and the finiteness of $L(\beta) = \tilde{L}^2(\beta)$ imply that $L^2(\beta^2) < \infty$.

For the purposes of a lower bound to be derived below (see (5.45) and Lemma 5.13), we must produce for both $i = 1$ and $i = 2$ a vector $\bar{\beta}^i$ that lies in the set $\mathrm{ri}(\mathrm{dom}\, L^i)$, the relative interior of the effective domain of L^i, and that satisfies $(\bar{\beta}^i)_1 > 0$. The existence of such a vector $\bar{\beta}^1$ is guaranteed by Hypothesis (H) and is proved in Lemma 4.4. The vector β^2 in the previous paragraph satisfies $L^2(\beta^2) < \infty$ and $(\beta^2)_1 > 0$. Hence there exists a vector $\bar{\beta}^2$ that lies in the set $\mathrm{ri}(\mathrm{dom}\, L^2)$ and that satisfies $(\bar{\beta}^2)_1 > 0$. Define real numbers

$$M = \max\{L^1(\bar{\beta}^1), L^2(\bar{\beta}^2)\} < \infty \quad \text{and} \quad \xi = \min\{(\bar{\beta}^1)_1, (\bar{\beta}^2)_1\} > 0$$

and let A be the convex hull of the set $\{\bar{\beta}^1, \bar{\beta}^2\}$.

We will use the lower bound

$$P\{S_n/n \in B(\beta, \varepsilon)\} \geqq P\{E_{1,n} \cap E_{2,n} \cap E_{3,n}\} , \tag{5.43}$$

where

$$E_{1,n} = \{S_{[n\gamma^0]}/n \in B(\gamma^0\beta^0, \delta_1)\} ,$$
$$E_{2,n} = \left\{S_{[n(\gamma^0 + \tau)]}/n \in \bigcup_{\beta' \in A} B(\gamma^0\beta^0 + \tau\beta', \delta_2)\right\} ,$$
$$E_{3,n} = \{S_n/n \in B(\beta, \varepsilon)\} .$$

In these definitions, $\delta_1 > 0$, $\delta_2 > 0$ and $\tau \in (0, 1 - \gamma^0)$ will be chosen below; the square brackets denote the integer part. By conditioning, the lower bound (5.43) may be rewritten in the form

$$P\{S_n/n \in B(\beta, \varepsilon)\} \geqq P_{1,n} \cdot P_{2,n} \cdot P_{3,n} ,$$

where

$$P_{1,n} = P\{E_{1,n}\} \quad \text{and} \quad P_{i,n} = P\{E_{i,n}|E_{i-1,n}\} \quad \text{for } i = 2, 3 .$$

We now obtain lower bounds on the limits inferior of $n^{-1} \log P_{1,n}$, $n^{-1} \log P_{2,n}$, and $n^{-1} \log P_{3,n}$.

For any $\delta_1 > 0$ our work in Subsection 5b implies that

$$\liminf_{n \to \infty} n^{-1} \log P_{1,n} \geqq -\gamma^0 \tilde{L}^0(\beta^0) . \tag{5.44}$$

If δ_1 is small compared to δ_2, then the occurrence of $E_{2,n}$ given $E_{1,n}$ is implied by S_n "tracking" $\bar{\beta}^1$ while in the halfspace Λ^1 and "tracking" $\bar{\beta}^2$ while in the halfspace Λ^2. Since $M \doteq \max\{L^1(\bar{\beta}^1), L^2(\bar{\beta}^2)\} < \infty$, this tracking should be possible at a cost that is less than M times the interval of time over which the tracking occurs. In fact, whenever $\delta_2 > \delta_1$, we have the following estimate proved in the next lemma:

$$\liminf_{n \to \infty} n^{-1} \log P_{2,n} \geqq -\tau M . \tag{5.45}$$

Lemma 5.13 *For any $\delta > 0$ and any $x \in \mathbb{R}^d$,*

$$\liminf_{n \to \infty} n^{-1} \log P_x\left\{(S_n - x)/n \in \bigcup_{\beta' \in A} B(\beta', \delta)\right\} \geqq -\max\{L^1(\bar{\beta}^1), L^2(\bar{\beta}^2)\} ,$$

where P_x denotes probability conditioned on $S_0 = x$.

Proof. The proof uses the same ideas as those of Subsection 5b, but in a much simpler way. Hence, the proof will only be sketched. Since for $i = 1, 2$ the vectors $\bar{\beta}^i$ lie in the sets ri(dom L^i), there exist vectors α^1 and α^2 in $\mathbb{R}^d$ (see the proof of Lemma 5.8(b)) satisfying the equalities

$$\nabla H^i(\alpha^i) = \bar{\beta}^i \quad \text{and} \quad L^i(\bar{\beta}^i) = \langle \alpha^i, \bar{\beta}^i \rangle - H^i(\alpha^i) .$$

For $x \in \mathbb{R}^d$ and $n \in \mathbb{N}$, we then define probability measures $P_{x,n}$ and $P_{x,n}^{\alpha^1, \alpha^2}$ analogously to the probability measures P_n and $P_n^{\alpha^1, \alpha^2}$ in Subsection 5b. For $i = 1, 2$ and $n \in \mathbb{N}$, we also define quantities v_n^i and r_n^i in Eq. (5.2). Then, as in Lemma 5.9(b), we have for $i = 1, 2$

$$v_n^i - r_n^i \bar{\beta}^i \to 0 \text{ in probability with respect to } P_{x,n}^{\alpha^1, \alpha^2} \text{ as } n \to \infty .$$

Since

$$P_x\left\{(S_n - x)/n \in \bigcup_{\beta' \in A} B(\beta', \delta)\right\} \geqq P_{x,n}\{\|v_n^i - r_n^i \bar{\beta}^i\| \leqq \delta/2, \quad i = 1, 2\}$$

$$= \int_{\cap_{i=1,2}\{\|v_n^i - r_n^i \bar{\beta}^i\| \leqq \delta/2,\, i=1,2\}} \frac{dP_{x,n}}{dP_{x,n}^{\alpha^1, \alpha^2}} dP_{x,n}^{\alpha^1, \alpha^2}$$

and since $\log(dP_{x,n}/dP_{x,n}^{\alpha^1,\alpha^2})$ has the form indicated in (5.20), we have the lower bound

$$\liminf_{n\to\infty} n^{-1}\log P_x\left\{(S_n - x)/n \in \bigcup_{\beta'\in A} B(\beta',\delta)\right\}$$
$$\geqq -\max\{L^1(\bar\beta^1), L^2(\bar\beta^2)\} - \gamma(\delta),$$

where $\gamma(\delta)\to 0$ as $\delta\to 0$. This yields the lemma. □

In order to complete the proof of the lower bound (5.41), we are left with estimating $\liminf_{n\to\infty} n^{-1}\log P_{3,n}$. In order to do this, we will use the following well-known process-level result for sums of i.i.d. random vectors. This result can be derived as a special case of Theorems 6.1′–6.2′ in A.D. Wentzell [31].

Lemma 5.14 *Let $\{X_i, i\in\mathbb{N}\}$ be a sequence of i.i.d. random vectors with a distribution given either by μ^1 or μ^2. Fix $T>0$ and take the function I to be either L^1 and L^2 depending on whether the X_i's have distribution μ^1 or μ^2, respectively. Then for any $\beta\in\mathbb{R}^d$ and any $\delta>0$*

$$\liminf_{n\to\infty} n^{-1}\log P\left\{\sup_{k\in\{1,\ldots,[Tn]\}}\left\|\frac{1}{n}\sum_{i=1}^{k}X_i - \frac{k}{n}\beta\right\|\leqq\delta\right\}\geqq -T\,I(\beta),$$

where the square brackets denote the integer part.

For ω in the conditioning set $E_{2,n}$ we have that

$$S_{[n(\gamma^0+\tau)]}/n\in\bigcup_{\beta'\in A} B(\gamma^0\beta^0+\tau\beta',\delta_2).$$

If $\delta_2\in(0,\xi\tau/2)$, then the set $\bigcup_{\beta'\in A}B(\gamma^0\beta^0+\tau\beta',\delta_2)$ is at least a distance $\xi\tau/2>0$ away from the boundary $\{x\in\mathbb{R}^d : x_1 = 0\}$. Defining the quantity

$$p_n = P\left\{\sup_{k\in\{[(\gamma^0+\tau)n]+1,\ldots,n\}}\left\|\frac{1}{n}\sum_{i=[(\gamma^0+\tau)n]+1}^{k} X_i^2 - \left(\frac{k-[(\gamma^0+\tau)n]}{n}\right)\beta^2\right\|\right.$$
$$\left.\leqq \xi\tau/4\right\},$$

we have for $\xi\tau\in(0,\varepsilon/4)$ the lower bound

$$P_{3,n}\geqq p_n\,. \qquad (5.46)$$

The latter is a consequence of the Markov property and the fact that for ω in the set $E_{2,n}\cap E_{3,n}$ the random walk remains in the right halfspace, in which it is spatially homogeneous. Using Lemma 5.14, we have

$$(5.47)\qquad \liminf_{n\to\infty} n^{-1}\log P_{3,n} \geqq \liminf_{n\to\infty} n^{-1}\log p_n \geqq -(1-\gamma^0-\tau)L^2(\beta^2) \geqq -\gamma^2 L^2(\beta^2)\,.$$

Combining (5.44), (5.45), and (5.47), we obtain

$$(5.48)\qquad \liminf_{n\to\infty} n^{-1}\log P\{S_n/n \in B(\beta,\varepsilon)\} \geqq -\gamma^0\tilde{L}^0(\beta^0) - \tau M - \gamma^2 L^2(\beta^2)\,.$$

Given $\varepsilon > 0$ and $\gamma > 0$, we assume that the following choices have been made:

(a) $\tau \in (0, \min\{1-\gamma^0, \varepsilon/4\xi, \gamma/2M\})$ so that (5.46) and (5.47) hold.
(b) $\delta_2 \in (0, \xi\tau/2)$ so that (5.46) and (5.47) hold.
(c) $\delta_1 \in (0, \delta_2)$ so that (5.45) holds.

For these choices, (5.42) and (5.48) imply that

$$\liminf_{n\to\infty} n^{-1}\log P\{S_n/n \in B(\beta,\varepsilon)\} \geqq -\tilde{L}^2(\beta) - \gamma = -L(\beta) - \gamma\,.$$

This yields the lower bound (5.41) since $\varepsilon > 0$ and $\gamma > 0$ are arbitrary.

The proofs of the large deviation bounds in Theorem 2.1 are now complete.

Acknowledgements. We are greatly indebted to an anonymous referee for suggestions on shortening the proof in Subsection 5a. The large deviation problem for the random vectors defined in (1.1) was suggested to us by Armand Makowsky. Jonathan Borwein pointed out Proposition 4.2. Richard S. Ellis thanks the Faculty of Industrial Engineering and Management at the Technion for their hospitality during his visit in the spring semester of 1989. He also thanks Kenji De Lige for Kenji's generous sharing of his LATEX-pertise. The figures in Sect. 3 were generated using Mathematica with the help of Jonathan Machta.

References

1. Azencott, R., Ruget, G.: Mélanges d'équations differentielles et grand écart à la loi des grandes nombres. Z. Wahrscheinlichkeitstheor. Verw. Geb. **38**, 1–54 (1977)
2. Bolthausen, E.: Markov process large deviations in τ-topology. Stochastic Processes Appl. **25**, 95–108 (1987)
3. Cramér, H.: Sur un nouveau théorème-limite de la théorie des probabilités. Colloque consacré à la théorie des probabilités, vol. 3. Paris: Hermann. Actual. Sci. Ind. **736**, 2–23 (1938)
4. de Acosta, A.: Upper bounds for large deviations of dependent random vectors. Z. Wahrscheinlichkeitstheor. Verw. Geb. **69**, 551–565 (1985)
5. de Acosta, A.: Large deviations for vector-valued functionals of a Markov chain: lower bounds. Ann. Probab. **16**, 925–960 (1988)
6. de Acosta, A.: Large deviations for empirical measures of Markov chains. J. Theoret. Probab. **3**, (1990) 395–431
7. Deuschel, J.-D., Stroock, D.W.: Large deviations. New York: Academic Press 1989
8. Donsker, M.D., Varadhan, S.R.S.: Asymptotic evaluation of certain Markov process expectations for large time, I. Commun. Pure Appl. Math. **28**, 1–47 (1975)
9. Donsker, M.D., Varadhan, S.R.S.: Asymptotic evaluation of certain Markov process expectations for large time, II. Commun. Pure Appl. Math. **28**, 279–301 (1975)

10. Donsker, M.D., Varadhan, S.R.S.: Asymptotic evaluation of certain Markov process expectations for large time, III. Commun. Pure Appl. Math. **29**, 389–461 (1976)
11. Donsker, M.D., Varadhan, S.R.S.: Asymptotic evaluation of certain Markov process expectations for large time, IV. Commun. Pure Appl. Math. **36**, 183–212 (1983)
12. Dupuis, P., Ellis, R.S., Weiss, A.: Large deviations for Markov processes with discontinuous statistics, I: General upper bounds. 1989 Ann. Probab. **9** 1280–1297 (1991)
13. Dupuis, P., Ishii, H., Soner, H.M.: A viscosity solution approach to the asmptotic analysis of queueing systems. Ann. Probab. **18**, 226–255 (1990)
14. Ellis, R.S.: Large deviations for a general class of random vectors. Ann. Probab. **12**, 1–12 (1984)
15. Ellis, R.S.: Entropy, large deviations and statistical mechnics. Berlin Heidelberg New York: Springer 1985
16. Ellis, R.S.: Large deviations for the empirical measure of a Markov chain with an application to the multivariate empirical measure. Ann. Probab. **16**, 1496–1508 (1988)
17. Ellis, R.S., Wyner, A.D.: Uniform large deviation property of the empirical process of a Markov chain. Ann. Probab. **17**, 1147–1151 (1989)
18. Freidlin, M.I., Wentzell, A.D.: Random perturbations of dynamical systems. Berlin Heidelberg New York: Springer 1984
19. Gärtner, J.: On large deviations from the invariant measure. Theory Probab. Appl. **22**, 24–39 (1977)
20. Iscoe, I., Ney, P., Nummelin, E.: Large deviations of uniformly recurrent Markov additive processes. Adv. Appl. Math. **6**, 373–412 (1985)
21. Jain, N.: Large deviation lower bounds for additive functionals of Markov processes. Ann. Probab. **18**, 1071–1098 (1990)
22. Korostelev, A.P., Leonov, S.L.: Action functional for a diffusion in discontinuous media. Technical report, Institute for Systems Studies, USSR Academy of Sciences, 1991
23. Korostelev, A.P., Leonov, S.L.: Action functional for diffusion process with discontiuous drift. Technical report, Institute for Systems Studies, USSR Academy of Sciences, 1991. Theory Probab. Appl. to appear
24. Kushner, H.: Introduction to stochastic control. New York: Holt, Rinehart and Winston 1970
25. Kushner, H.J.: Approximation and weak convergence methods for random processes with applications to stochastic system theory. Cambridge: MIT Press 1984
26. Kushner, H.J.: Weak convergence methods and singularly perturbed stochastic control and filtering problems. Boston: Birkhäuser 1990
27. Parekh, S., Walrand, J.: Quick simulation method for excessive backlogs in networks of queues. In: Lions, P.-L., Fleming, W., (eds.) Proceedings of the IMA, 1986
28. Rockafellar, R.T.: Convex analysis. Princeton: Princeton University 1970
29. Stroock, D.W.: An introduction to the theory of large deviations. Berlin Heidelberg New York: Springer 1984
30. Wentzell, A.D.: Rough limit theorems on large deviations for Markov stochastic processes, I. Theory Probab. Appl. **21**, 227–242 (1976)
31. Wentzell, A.D.: Rough limit theorems on large deviations for Markov stochastic processes, II. Theory Probab. Appl. **21**, 499–512 (1976)
32. Wentzell, A.D.: Rough limit theorems on large deviations for Markov stochastic processes, III. Theory Probab. Appl. **24**, 675–692 (1979)
33. Wentzell, A.D.: Rough limit theorems on large deviations for Markov stochastic processes, IV. Theory Probab. Appl. **27**, 215–234 (1982)