DOI: 10.1214/ECP.v1-973

Elect. Comm. in Probab. **1** (1996) 11–17

ELECTRONIC COMMUNICATIONS in PROBABILITY

MODERATE DEVIATIONS FOR MARTINGALES WITH BOUNDED JUMPS [1]

A. DEMBO[2]
Department of Electrical Engineering, Technion Israel Institute of Technology, Haifa 32000, ISRAEL.
e-mail: amir@playfair.stanford.edu

Submitted: 3 September 1995; Revised: 25 February 1996.

AMS 1991 Subject classification: 60F10,60G44,60G42
Keywords and phrases: Moderate deviations, martingales, bounded martingale differences.

Abstract
We prove that the Moderate Deviation Principle (MDP) holds for the trajectory of a locally square integrable martingale with bounded jumps as soon as its quadratic covariation, properly scaled, converges in probability at an exponential rate. A consequence of this MDP is the tightness of the method of bounded martingale differences in the regime of moderate deviations.

1 Introduction

Suppose $\{X_m, \mathcal{F}_m\}_{m=0}^{\infty}$ is a discrete-parameter real valued martingale with bounded jumps $|X_m - X_{m-1}| \le a$, $m \in \mathbb{N}$, filtration $\mathcal{F}_m$ and such that $X_0 = 0$. The basic inequality for the method of bounded martingale differences is Azuma-Hoeffding inequality (c.f. [1]):

$$\mathbb{P}\{X_k \ge x\} \le e^{-x^2/2ka^2} \qquad \forall x > 0. \tag{1}$$

In the special case of i.i.d. differences $\mathbb{P}\{X_m - X_{m-1} = a\} = 1 - \mathbb{P}\{X_m - X_{m-1} = -\epsilon a/(1-\epsilon)\} = \epsilon \in (0,1)$, it is easy to see that $\mathbb{P}\{X_k \ge x\} \le \exp[-kH(\epsilon + (1-\epsilon)x/(ak)|\epsilon)]$, where $H(q|p) = q\log(q/p) + (1-q)\log((1-q)/(1-p))$. For $\epsilon \to 0$, the latter upper bound approaches 0, thus demonstrating that (1) may in general be a non-tight upper bound. Let $B(u) = 2u^{-2}((1+u)\log(1+u) - u)$ and

$$\langle X \rangle_m = \sum_{k=1}^{m} E[(X_k - X_{k-1})^2 | \mathcal{F}_{k-1}]$$

denote the quadratic variation of $\{X_m, \mathcal{F}_m\}_{m=0}^{\infty}$. Then,

$$\mathbb{P}\{X_k \ge x\} \le \mathbb{P}\{\langle X \rangle_k \ge y\} + e^{-x^2 B(ax/y)/2y} \qquad \forall x, y > 0 \tag{2}$$

[1] Partially supported by NSF DMS-9209712 and DMS-9403553 grants and by a US-ISRAEL BSF grant
[2] On leave from the Department of Mathematics and Department of Statistics, Stanford University, Stanford, CA 94305

(c.f. [4, Theorem (1.6)]). In particular, $B(0_+) = 1$, recovering (1) for the choice $y = ka^2$ and $x/y \to 0$. The inequality (2) holds also for the more general setting of locally square integrable (continuous-parameter) martingales with bounded jumps (c.f. [7, Theorem II.4.5]).

In this note we adopt the latter setting and demonstrate the tightness of (2) in the range of moderate deviations, corresponding to $x/y \to 0$ while $x^2/y \to \infty$ (c.f. Remark 5 below). We note in passing that for *continuous* martingales [6] studies the tightness of the inequality:

$$\mathbb{P}\{X_k \geq \frac{1}{2}\, x(1 + \langle X\rangle_k/y)\} \leq e^{-x^2/2y} \,,$$

using Girsanov transformations, whereas we apply large deviation theory and concentrate on martingales with (bounded) jumps, encompassing the case of discrete-parameter martingales.

Recall that a family of random variables $\{Z_k; k > 0\}$ with values in a topological vector space $\mathcal{X}$ equipped with σ-field $\mathcal{B}$ satisfies the Large Deviation Principle (LDP) with speed $a_k \downarrow 0$ and *good rate function* $I(\cdot)$ if the level sets $\{x; I(x) \leq \alpha\}$ are compact for all $\alpha < \infty$ and for all $\Gamma \in \mathcal{B}$

$$-\inf_{x\in\Gamma^o} I(x) \;\leq \liminf_{k\to\infty} a_k \log \mathbb{P}\{Z_k \in \Gamma\} \leq \limsup_{k\to\infty} a_k \log \mathbb{P}\{Z_k \in \Gamma\} \leq -\inf_{x\in\overline{\Gamma}} I(x)$$

(where Γ^o and $\overline{\Gamma}$ denote the interior and closure of Γ, respectively). The family of random variables $\{Z_k; k > 0\}$ satisfies the *Moderate Deviation Principle* with good rate function $I(\cdot)$ and critical speed $1/h_k$ if for every speed $a_k \downarrow 0$ such that $h_k a_k \to \infty$, the random variables $\sqrt{a_k} Z_k$ satisfy the LDP with the good rate function $I(\cdot)$.

Let $D(\mathbb{R}^d)(= D(\mathbb{R}_+, \mathbb{R}^d))$ denote the space of all $\mathbb{R}^d$-valued *càdlàg* (i.e. right-continuous with left-hand limits) functions on $\mathbb{R}_+$ equipped with the locally uniform topology. Also, $C(\mathbb{R}^d)$ is the subspace of $D(\mathbb{R}^d)$ consisting of continuous functions.

The process $X \in D(\mathbb{R}^d)$ is defined on a complete stochastic basis $(\Omega, \mathcal{F}, \mathbf{F} = \mathcal{F}_t, \mathbb{P})$ (c.f. [5, Chapters I and II] or [7, Chapters 1-4] for this and the related definitions that follow). We equip $D(\mathbb{R}^d)$ hereafter with a σ-field $\mathcal{B}$ such that $X : \Omega \to D(\mathbb{R}^d)$ is measurable ($\mathcal{B}$ may well be strictly smaller than the Borel σ-field of $D(\mathbb{R}^d)$).

Suppose that $X \in \mathcal{M}^2_{\mathrm{loc},0}$ is a locally square integrable martingale with bounded jumps $|\Delta X| \leq a$ (and $X_0 = 0$). We denote by (A, C, ν) the triplet predictable characteristics of X, where here $A = 0$, $C = (C_t)_{t\geq 0}$ is the $\mathbf{F}$-predictable quadratic variation process of the continuous part of X and $\nu = \nu(ds,\, dx)$ is the $\mathbf{F}$-compensator of the measure of jumps of X. Without loss of generality we may assume that

$$\nu(\{t\}, \mathbb{R}^d) = \int_{|x|\leq a} \nu(\{t\},\, dx) \leq 1, \quad \int_{|x|\leq a} x\nu(\{t\}, dx) = 0, \quad t > 0 \tag{3}$$

and for all $s < t$, $(C_t - C_s)$ is a symmetric positive-semi-definite $d \times d$ matrix. The predictable quadratic characteristic (covariation) of X is the process

$$\langle X\rangle_t = C_t + \int_0^t \int_{|x|\leq a} xx' d\nu \,, \tag{4}$$

where x' denotes the transpose of $x \in \mathbb{R}^d$, and $\|A\| = \sup_{|\lambda|=1} |\lambda' A\lambda|$ for any $d \times d$ symmetric matrix A.

Our main result is as follows.

Proposition 1 *Suppose the symmetric positive-semi-definite $d \times d$ matrix Q and the regularly varying function h_t of index $\alpha > 0$ are such that for all $\delta > 0$:*

$$\limsup_{t\to\infty} h_t^{-1} \log \mathbb{P}\{\|h_t^{-1}\langle X\rangle_t - Q\| > \delta\} < 0 \ . \tag{5}$$

Then $\left\{h_k^{-1/2} X_{k\cdot}\right\}$ satisfies the MDP in $(D[\mathbb{R}^d], \mathcal{B})$ (equipped with the locally uniform topology) with critical speed $1/h_k$ and the good rate function

$$I(\phi) = \begin{cases} \int_0^\infty \Lambda^*(\dot\phi(t))\alpha^{-1}t^{(1-\alpha)}dt & \phi \in \mathcal{AC}_0 \\ \infty & \text{otherwise,} \end{cases} \tag{6}$$

where $\Lambda^*(v) = \sup_{\lambda\in\mathbb{R}^d}(\lambda' v - \frac{1}{2}\lambda' Q\lambda)$, and $\mathcal{AC}_0 = \{\phi : \mathbb{R}_+ \to \mathbb{R}^d$ with $\phi(0) = 0$ and absolutely continuous coordinates $\}$.

Remark 1 Note that both (5) and the MDP are invariant to replacing h_t by g_t such that $h_t/g_t \to c \in (0, \infty)$ and taking cQ instead of Q. Thus, if $Q \neq 0$ we may take $h_t =$ median $\| \langle X\rangle_t \|$, and in general we may assume with no loss of generality that $h_t \in D(\mathbb{R}_+)$ is strictly increasing of bounded jumps.
Remark 2 If X is a locally square integrable martingale with independent increments, then $\langle X\rangle$ is a deterministic process, hence suffices that $h_t^{-1}\langle X\rangle_t \to Q$ for (5) to hold.

As stated in the next corollary, less is needed if only X_k (or $\sup_{s\le k} X_s$) is of interest.

Corollary 1
(a) Suppose that (5) holds for some unbounded h_t (possibly not regularly varying). Then, $\left\{h_k^{-1/2} X_k\right\}$ satisfies the MDP in $\mathbb{R}^d$ with critical speed $1/h_k$ and good rate function $\Lambda^*(\cdot)$.
(b) If also $d = 1$, then $\left\{h_k^{-1/2} \sup_{s\le k} X_s\right\}$ satisfies the MDP with the good rate function $I(z) = z^2/(2Q)$ for $z \ge 0$ and $I(z) = \infty$ otherwise.

Remark 3 For $d = 1$, discrete-time martingales, and assuming that $h_k = \langle X\rangle_k$ is *non-random*, strong Normal approximation for the law of $h_k^{-1/2} X_k$ is proved in [9] for the range of values corresponding to $a_k^3 h_k \to \infty$.
Remark 4 The difference between Proposition 1 and Corollary 1 is best demonstrated when considering $X_t = B_{h_t}$, with B_s the standard Brownian motion. The MDP for $h_t^{-1/2} B_{h_t}$ in $\mathbb{R}$ then trivially holds, whereas the MDP for $h_k^{-1/2} B_{h_{tk}}$ is equivalent to Schilder's theorem (c.f. [3, Theorem 5.2.3]), and thus holds only when h_t is regularly varying of index $\alpha > 0$.
Remark 5 When $d = 1$ and $Q \neq 0$, the rate function for the MDP of part (a) of Corollary 1 is $x^2/(2Q)$. For $y = h_k Q(1 + \delta)$, $\delta > 0$ and $x = x_k = o(y)$ such that $x^2/y \to \infty$, this MDP then implies that $\mathbb{P}\{X_k \ge x\} = \exp(-(1 + \delta + o(1))x^2/2y)$ while $P(\langle X\rangle_k \ge y) = o(\exp(-x^2/2y))$ by (5). Consequently, for such values of x, y the inequality (2) is tight for $k \to \infty$ (see also Remark 9 below for non-asymptotic results).
Remark 6 In contrast with Corollary 1 we note that the LDP with speed m^{-1} may fail for $m^{-1}X_m$ even when X is a real valued discrete-parameter martingale with bounded independent increments such that $\langle X\rangle_m = m$. Specifically, let $b : \mathbb{N} \to \{1, 2\}$ be a deterministic sequence such that $p_m = m^{-1} \sum_{k=1}^m 1_{\{b(k)=1\}}$ fails to converge for $m \to \infty$ and let μ_i, $i = 1, 2$ be two probability measures on $[-a, a]$ such that $\int x d\mu_i = 0$, $\int x^2 d\mu_i = 1$, $i = 1, 2$ while $c_1 \neq c_2$ for

$c_i = \log \int e^x d\mu_i$. Then, ΔX_k independent random variables of law $\mu_{b(k)}$, $k \in \mathbb{N}$, result with X_m as above. Indeed, $m^{-1} \log \mathbb{E}\{\exp(X_m)\} = p_m c_1 + (1-p_m)c_2$ fails to converge for $m \to \infty$, hence by Varadhan's lemma (c.f. [3, Theorem 4.3.1]), necessarily the LDP with speed m^{-1} fails for $m^{-1}X_m$.

Remark 7 Corollary 1 may fail when X is a real valued discrete-parameter martingale with *unbounded* independent increments such that $\langle X \rangle_m = m$. Specifically, for $m_j = 2^{2j^2}$, $j \in \mathbb{N}$ let $M(m_j) = 2(m_j \log m_j)^{1/2}$ and $M(k) = 1$ for all other $k \in \mathbb{N}$. Let Z_k be independent Bernoulli$(1/(M(k)^2+1))$ random variables. Then, $\Delta X_k = M(k)Z_k - M(k)^{-1}(1-Z_k)$ result with X_m as above, with the LDP of speed $1/\log m$ not holding for $(m \log m)^{-1/2} X_m$. Indeed, let Y_m be the martingale with ΔY_{m_j} i.i.d. and independent of X such that $\mathbb{P}\{\Delta Y_{m_j} = 1\} = \mathbb{P}\{\Delta Y_{m_j} = -1\} = 0.5$ and $\Delta Y_k = \Delta X_k$ for all other $k \in \mathbb{N}$. Then, $(m \log m)^{-1/2}|X_m - Y_m| \to 0$ for $m = (m_j - 1)$, $j \to \infty$, while $(m \log m)^{-1/2}(X_m - Y_m) \geq 2Z_m + o(1)$ for $m = m_j$, $j \to \infty$. The LDP with speed $1/\log m$ and good rate function $x^2/2$ holds for $(m \log m)^{-1/2} Y_m$ (c.f. Corollary 1), while $\log \mathbb{P}\{Z_{m_j} = 1\}/\log m_j \to -1$ as $j \to \infty$. Consequently, the LDP bounds fail for $\{(m \log m)^{-1/2} X_m \geq 2\}$.

Proposition 1 is proved in the next section with the proof of Corollary 1 provided in Section 3. Both results build upon Lemma 1. Indeed, Proposition 1 is a direct consequence of Lemma 1 and [8]. Also, with Lemma 1 holding, it is not hard to prove part (a) of Corollary 1 as a consequence of the Gärtner–Ellis theorem (c.f. [3, Theorem 2.3.6]), without relying on [8].

2 Proof of Proposition 1

The cumulant $G(\lambda) = (G_t(\lambda))_{t \geq 0}$ associated with X is

$$G_t(\lambda) = \frac{1}{2}\, \lambda' C_t \lambda + \int_0^t \int_{|x| \leq a} (e^{\lambda' x} - 1 - \lambda' x) \nu(ds, dx), t > 0, \lambda \in \mathbb{R}^d . \tag{7}$$

The stochastic (or the Doléans-Dade) exponential of $G(\lambda)$, denoted $\mathcal{E}(G(\lambda))$ is given by

$$\varphi_t(\lambda) = \log \mathcal{E}(G(\lambda))_t = G_t(\lambda) + \sum_{s \leq t} \left[\log(1 + \Delta G_s(\lambda)) - \Delta G_s(\lambda) \right] , \tag{8}$$

where

$$\Delta G_s(\lambda) = \int_{|x| \leq a} (e^{\lambda' x} - 1) \nu(\{s\}, dx) = \int_{|x| \leq a} (e^{\lambda' x} - 1 - \lambda' x) \nu(\{s\}, dx) . \tag{9}$$

The next lemma which is of independent interest, is key to the proof of Proposition 1.

Lemma 1 *For $\epsilon > 0$, let $v(\epsilon) = 2(e^\epsilon - 1 - \epsilon)/\epsilon^2 \geq 1 \geq v(-\epsilon) - \epsilon^2 v(\epsilon)^2/4 = w(\epsilon)$. Then, for any $0 \leq u \leq t < \infty$, $\quad \lambda \in \mathbb{R}^d$*

$$\frac{1}{2}\, w(|\lambda| a) \lambda' (\langle X \rangle_t - \langle X \rangle_u) \lambda \leq \varphi_t(\lambda) - \varphi_u(\lambda) \leq \frac{1}{2}\, v(|\lambda| a)\ \lambda' (\langle X \rangle_t - \langle X \rangle_u) \lambda. \tag{10}$$

Remark 8 Since $\exp[\lambda' X_t - \varphi_t(\lambda)]$ is a local martingale (c.f. [7, Section 4.13]), Lemma 1 implies that $\exp[\lambda' X_t - \frac{1}{2} v(|\lambda| a) \lambda' \langle X \rangle_t \lambda]$ is a non-negative super-martingale while $\exp[\lambda' X_t - \frac{1}{2} w(|\lambda| a) \lambda' \langle X \rangle_t \lambda]$ is a non-negative local sub-martingale. Noting that $w(|\lambda| a), v(|\lambda| a) \to 1$ for $|\lambda| \to 0$, these are to be compared with the local martingale property of $\exp[\lambda' X_t - \frac{1}{2} \lambda' \langle X \rangle_t \lambda]$ when $X \in \mathcal{M}^c_{\mathrm{loc},0}$ is a *continuous* local martingale (c.f. [7, Section 4.13]).

Remark 9 For $d = 1$ it follows that for every $\lambda \in \mathbb{R}$,

$$\mathbb{E}\{\exp[\lambda X_m - \frac{1}{2}v(|\lambda|a)\lambda^2\langle X\rangle_m]\} \leq 1 \tag{11}$$

(c.f. Remark 8). The inequality (2) then follows by Chebycheff's inequality and optimization over $\lambda \geq 0$. For the special case of a real-valued discrete-parameter martingale X_m also

$$\mathbb{E}\{\exp[\lambda X_m - \frac{1}{2}w(|\lambda|a)\lambda^2\langle X\rangle_m]\} \geq 1\,, \tag{12}$$

and we can even replace $w(|\lambda|a)$ in (12) by $v(-|\lambda|a)$ (c.f. [4, (1.4)] where the sub-martingale property of $\exp(\lambda X_m - \frac{1}{2}v(-|\lambda|a)\lambda^2\langle X\rangle_m)$ is proved).

Proof: To prove the upper bound on $\varphi_t(\lambda) - \varphi_u(\lambda)$ note that $\log(1+x) - x \leq 0$ implying by (8) that $\varphi_t(\lambda) - \varphi_u(\lambda) \leq G_t(\lambda) - G_u(\lambda)$. The required bound then follows from (7) since $(e^{\lambda' x} - 1 - \lambda' x) \leq \frac{1}{2}\, v(|\lambda|a)\lambda'(xx')\lambda$ for $|x| \leq a$, and $\lambda'(C_t - C_u)\lambda \geq 0$ for $u \leq t$.
To establish the corresponding lower bound, note that since $\Delta G_s(\lambda) \geq 0$ (see (9)) and $\log(1+x) - x \geq -x^2/2$ for all $x \geq 0$, we have that

$$\varphi_t(\lambda) - \varphi_u(\lambda) \geq G_t(\lambda) - G_u(\lambda) - \frac{1}{2}\sum_{u<s\leq t}\Delta G_s(\lambda)^2\,.$$

Moreover, again by (9) we see that

$$0 \leq \Delta G_s(\lambda) \leq \frac{1}{2}\, v(|\lambda|a)\lambda'\left[\int_{|x|\leq a} xx'\nu(\{s\},dx)\right]\lambda \leq \frac{1}{2}\, v(|\lambda|a)^2(|\lambda|a)^2\,.$$

Hence,

$$\begin{aligned}\frac{1}{2}\sum_{u<s\leq t}\Delta G_s(\lambda)^2 &\leq \frac{1}{8}v(|\lambda|a)^2(|\lambda|a)^2\lambda'\left[\sum_{u<s\leq t}\int_{|x|\leq a} xx'\nu(\{s\},dx)\right]\lambda \\ &\leq \frac{1}{8}v(|\lambda|a)^2(|\lambda a)^2\lambda'\,[\langle X\rangle_t - \langle X\rangle_u]\,\lambda\,,\end{aligned}$$

and the required lower bound follows by noting that

$$G_t(\lambda) - G_u(\lambda) \geq \frac{1}{2}\, v(-|\lambda|a)\lambda'[\langle X\rangle_t - \langle X\rangle_u]\lambda\,. \qquad \blacksquare$$

To prove Proposition 1 we need the following immediate consequence of Lemma 1.

Lemma 2 *Suppose there exists $q \in C[0,\infty)$, a positive-semi-definite matrix Q and an unbounded function $h:\ \mathbb{R}_+ \to \mathbb{R}_+$ such that for all $\delta > 0, T < \infty$*

$$\limsup_{k\to\infty}\frac{1}{h_k}\log \mathbb{P}\left\{\sup_{u\in[0,T]}\left\|\frac{\langle X\rangle_{uk}}{h_k} - q(u)Q\right\| > \delta\right\} < 0\,. \tag{13}$$

Then, for every $\lambda \in \mathbb{R}^d$ and $a_k \to 0$ such that $h_k a_k \to \infty$,

$$\limsup_{k\to\infty} a_k \log \mathbb{P}\left\{\sup_{u\in[0,T]}\left|a_k\varphi_{uk}(\lambda/\sqrt{h_k a_k}) - \frac{1}{2}\,q(u)\lambda' Q\lambda\right| > \delta\right\} = -\infty\,. \tag{14}$$

Proof: Use (10), noting that $a_k = \frac{1}{h_k}(a_k h_k)$ with $a_k h_k \to \infty$, and that $\lim_{k\to\infty} v(|\lambda|a/\sqrt{a_k h_k}) = \lim_{k\to\infty} w(|\lambda|a/\sqrt{a_k h_k}) = 1$, while $\sup_{u\in[0,T]} |q(u)| < \infty$. ∎

The next lemma is a simple application of the results of [8], relating (14) with the LDP (with speed a_k) of $\left\{\sqrt{\frac{a_k}{h_k}} X_{k\cdot}\right\}$.

Lemma 3 *When (14) holds, the processes* $\left\{\sqrt{\frac{a_k}{h_k}} X_{k\cdot}, k>0\right\}$ *satisfy the LDP in* $(D(\mathbb{R}^d), \mathcal{B})$ *with speed* a_k *and the good rate function*

$$I(\phi) = \begin{cases} \int_0^\infty \Lambda^*\left(\frac{d\phi}{dq}(t)\right) q(dt) & \phi \ll q, \quad \phi(0) = 0 \\ \infty & \text{otherwise} \end{cases} \tag{15}$$

(where $q \in M_+(\mathbb{R}_+)$ *is the continuous locally finite measure on* $(\mathbb{R}_+, \mathcal{B}_{\mathbb{R}_+})$ *such that* $q([0,t]) = q(t)$*).*

Proof: For each sequence $k_n \to \infty$ we shall apply [8, Theorem 2.2] for the local martingales $\sqrt{a_{k_n}/h_{k_n}} X_{k_n t}$ replacing $\frac{1}{n}$ throughout by a_{k_n}. Cramér's condition [8, (2.6)] is trivially holding in the current setting, while for $G_t(\lambda) = \frac{1}{2} q(t)\lambda' Q\lambda$ the condition ($\sup \mathcal{E}$) of [8, Theorem 2.2] is merely (14). Moreover, for this $G_t(\lambda)$ the condition [8, (G)] is easily shown to hold (as $H_{s,t}(\cdot)$ is then a positive-definite quadratic form on the linear subspace dom$H_{s,t}$ for all $s < t$). Thus, the LDP in Skorohod topology follows from [8, Theorem 2.2] and the explicit form (15) of the rate function follows from [8, (2.4)] taking there $g_t(\lambda) = \frac{1}{2}\lambda' Q\lambda$. Suppose $I(\phi) < \infty$. Then, $\phi \ll q$ and since $q \in C[0,\infty)$ it follows that $\phi \in C(\mathbb{R}^d)$. Hence, by [8, Theorem C] we may replace the Skorohod topology by the stronger locally uniform topology on $D(\mathbb{R}^d)$. ∎

Proposition 1 follows by combining Lemmas 2 and 3 with the next lemma.

Lemma 4 *If* h_t *is regularly varying of index* $\alpha > 0$ *then (5) implies that (13) holds for* $q(u) = u^\alpha$.

Proof: Fix $T < \infty$ and $\delta > 0$. Since h_t is regularly varying of index $\alpha > 0$, clearly $h_{uk}/h_k \to u^\alpha$ for all $u \in (0,\infty)$ (c.f. [2, page 18]). Take $\epsilon > 0$ small enough for $\sup_{0\le i\le \lceil T/\epsilon\rceil} |q(i\epsilon+\epsilon) - q(i\epsilon)| \le \delta/(3\|Q\|)$, and $k_0 < \infty$ such that $\sup_{0\le i\le\lceil T/\epsilon\rceil} |h_{i\epsilon k}/h_k - q(i\epsilon)| \le \delta/(3\|Q\|)$ whenever $k \ge k_0$ (note that $q(0) = 0$).
The monotonicity of $\langle X\rangle_{tk}$ in t (and $\langle X\rangle_0 = 0$) implies that for all $k \ge k_0$

$$\left\{\sup_{u\in[0,T]} \left\|\frac{\langle X\rangle_{uk}}{h_k} - q(u)Q\right\| > \delta\right\} \subseteq \left\{\sup_{1\le i\le\lceil T/\epsilon\rceil} \|\langle X\rangle_{i\epsilon k} - h_{i\epsilon k}Q\| > \frac{1}{3}\delta h_k\right\}.$$

Hence, suffices to show that for every $i \in \mathbb{N}$, $\epsilon > 0$

$$\limsup_{k\to\infty} \frac{1}{h_k} \log \mathbb{P}\left\{\| \langle X\rangle_{i\epsilon k} - h_{i\epsilon k}Q \| > \frac{1}{3}\delta h_k\right\} < 0\,.$$

Since $h_{i\epsilon k}/h_k \to q(i\epsilon) \in (0,\infty)$ this inequality follows from (5). ∎

3 Proof of Corollary 1

(a) Assume first that h_t is regularly varying of index 1. Given Proposition 1, this case is easily settled by applying the contraction principle for the continuous mapping $\phi \mapsto \phi(1) : D[\mathbb{R}^d] \to \mathbb{R}^d$. In the general case, we take without loss of generality $h_t \in D(\mathbb{R}_+)$ strictly increasing of bounded jumps (see Remark 1). Let $\sigma_s = \inf\{t \geq 0 : h_t \geq s\}$ and $g_s = h_{\sigma_s}$. Note that $g_s - s$ is bounded, while (5) holds for the locally square integrable martingale $Y_s = X_{\sigma_s}$ of bounded jumps and the regularly varying function g_s of index 1. Consequently, $\{g_s^{-1/2} Y_s\}$ satisfies the MDP with the critical speed $1/g_s$ and the good rate function $\Lambda^*(\cdot)$. Since h_t is strictly increasing and unbounded it follows that $\sigma(\mathbb{R}_+) = \mathbb{R}_+$. Hence, this MDP is equivalent to the MDP for $\{h_k^{-1/2} X_k\}$.
(b) As in part (a) above suffices to prove the stated MDP for h_t regularly varying of index 1. Applying the contraction principle for the continuous mapping $\phi \mapsto \sup_{s\leq 1} \phi(s)$ we deduce the stated MDP from Proposition 1. Since $\Lambda^*(v) = v^2/(2Q)$, the good rate function for this MDP is (c.f. (6))

$$I(z) = \frac{1}{2Q} \inf_{\{\phi \in \mathcal{AC}_0 : \sup_{s\leq 1} \phi(s) = z\}} \int_0^\infty \dot{\phi}(s)^2 ds \geq \frac{z^2}{2Q} .$$

Clearly, $\phi(0) = 0$ implies that $I(z) = \infty$ for $z < 0$, while taking $\phi(s) = (s \wedge 1)z$ we conclude that $I(z) = z^2/(2Q)$ for $z \geq 0$.

References

[1] K. Azuma (1967): Weighted sums of certain dependent random variables, *Tohoku Math. J.* **3**, 357–367.

[2] N.H. Bingham, C.M. Goldie and J.L. Teugels (1987): *Regular Variation* Cambridge Univ. Press.

[3] A. Dembo and O. Zeitouni (1993): *Large Deviations Techniques and Applications* Jones and Bartlett, Boston.

[4] D. Freedman (1975): On tail probabilities for martingales, *Ann. Probab.* **3**, 100–118.

[5] J. Jacod and A.N. Shiryaev (1987): *Limit theorems for stochastic processes* Springer-Verlag, Berlin.

[6] D. Khoshnevisan (1995): Deviation inequalities for continuous martingales, (preprint)

[7] R. Sh.Liptser and A.N. Shiryaev (1989): *Theory of Martingales* Kluwer, Dorndrecht.

[8] A. Puhalskii (1994): The method of stochastic exponentials for large deviations, *Stoch. Proc. Appl.* **54**, 45–70.

[9] A. Rackauskas (1990): On probabilities of large deviations for martingales, *Liet. Matem. Rink.* **30**, 784–794.