# MODERATE DEVIATIONS FOR I.I.D. RANDOM VARIABLES [*]

## Peter Eichelsbacher[1] and Matthias Löwe[2]

**Abstract**. We derive necessary and sufficient conditions for a sum of i.i.d. random variables $\sum_{i=1}^{n} X_i/b_n$ – where $\frac{b_n}{n} \downarrow 0$, but $\frac{b_n}{\sqrt{n}} \uparrow \infty$ – to satisfy a moderate deviations principle. Moreover we show that this equivalence is a typical moderate deviations phenomenon. It is not true in a large deviations regime.

**Mathematics Subject Classification.** 60F10.

## 1. Introduction

Moderate Deviation Principles (MDPs for short) have become an increasingly popular topic in the recent probability literature. Formally an MDP for some random variable is just a large deviation principle. Here we say that a sequence of random variables $(Y_n)$ with topological state space $X$ obeys a large deviation principle with speed $a_n$ and good rate function $I(\cdot) : X \to \mathbb{R}_0^+$ if

- $I$ is lower semi-continuous and has compact level sets $N_L := \{x \in X : I(x) \leq L\}$, for every $L \in \mathbb{R}_0^+$;
- for every open set $G \subseteq X$ it holds

$$\liminf_{n \to \infty} \frac{1}{a_n} \log \mathbb{P}(Y_n \in G) \geq - \inf_{x \in G} I(x); \tag{1.1}$$

- for every closed set $A \subseteq X$ it holds

$$\limsup_{n \to \infty} \frac{1}{a_n} \log \mathbb{P}(Y_n \in A) \leq - \inf_{x \in A} I(x). \tag{1.2}$$

Formally there is no distinction between an MDP and a large deviation principle. Usually a large deviation principle lives on the scale of a law of large number type ergodic phenomenon, while MDPs describe the probabilities on a scale between a law of large numbers and some sort of central limit theorem. For both, large deviation principles and MDPs the three points listed under the bullets above serve as a definition. However, there are important differences in the behavior of the two principles. Typically, the rate function in a large deviation principle will depend on the distribution of the underlying random variables, while an MDP inherits

[1] Ruhr-Universität Bochum, Fakultät für Mathematik, NA3/68, 44780 Bochum, Germany;
e-mail: `peter.eichelsbacher@ruhr-uni-bochum.de`

[2] Department of Mathematics, University of Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, The Netherlands;
e-mail: `loewe@sci.kun.nl`

properties of both the central limit behavior as well as of the large deviation principle. For example, one often sees the exponential decay of moderate deviation probabilities which is typical of the large deviations. On the other hand the rate function in an MDP quite often is "non-parametric" in the sense that it only depends on the limiting density in the central limit theorem for these variables but not on individual characteristics of their distributions. Often even the rate function of an MDP interpolates between the logarithmic probabilities that can be expected from the central limit theorem and the large deviations rate function – even if the limit is not normal (see *e.g.* [7, 11]); situations where this is not the case are particularly interesting [3].

All of this also holds true in the case we are going to study in this note, the case of sums of $\mathbb{R}$-valued, i.i.d. random variables $X_1, X_2, \ldots$ Here the large deviation principle for the average has been known since the times of Cramér [4], also see [5]: if the log of the moment generating function $\Lambda(t)$ of one of the variables is finite in a neighborhood of zero, $\frac{1}{n}\sum_{i=1}^n X_i$ satisfies a large deviations principle with speed $n$ and rate function $\Lambda^*(\cdot)$, the Legendre transform of $\Lambda$. MDPs for $\frac{1}{b_n}\sum_{i=1}^n (X_i - \mathbb{E}X_i)$, where $\frac{b_n}{n} \downarrow 0$ but $\frac{b_n}{\sqrt{n}} \uparrow \infty$, have been studied by de Acosta [1]. Here – for all such sequences $(b_n)_n$ – the rate functions are given by $I(t) = \frac{1}{2\operatorname{Var}X_1}t^2$ and the speed is $a_n = b_n^2/n$. In a recent paper [2] Arcones states that assumption (2.3) below is necessary and sufficient for $\frac{1}{b_n}\sum_{i=1}^n (X_i - \mathbb{E}X_i)$ to satisfy a moderate deviations principle. As a matter of fact it seems that the search for such an equivalent condition is almost as old as the investigation of moderate deviations for means of i.i.d. random variables. Indeed, already Ledoux [9] discusses the connection of our assumption (2.3) below with the moderate deviations *upper* bound for a restricted range of normalizations $(b_n)$. Unfortunately, Arcones proof in [2] of this interesting equivalence is spread over several theorems and articles and technically very involved, since it relies on MDPs for Banach space valued stochastic processes. Moreover it is most unfortunately spoiled by several flaws and therefore hardly penetrable.

In this note we aim at giving a *self-contained and short* proof of Arcones's result. We briefly describe what assumption (2.3) means for a distribution by describing the borderline cases where it is not fulfilled or only fulfilled for some choices of the normalizing sequence $(b_n)$. Moreover we show that Theorem 2.2 below is a typical moderate deviation phenomenon. On the large deviation scale Theorem 2.2 does not hold true.

The techniques used are among others a truncation argument, the Gärtner–Ellis theorem, and Lévy's inequality. These and similar techniques have been employed extensively by Ledoux [9] and Djellout [6]. We will comment on Ledoux's result and techniques a bit more extensively in Remark 2.8.

## 2. The main result

In this section we state and prove the main result of this note. Let $X_1, X_2, \ldots$ be i.i.d. real-valued random variables and $(b_n)$ be an increasing sequence of numbers with

$$\frac{b_n}{\sqrt{n}} \uparrow \infty \quad \text{and} \quad \frac{b_n}{n} \downarrow 0. \tag{2.1}$$

The following theorem gives necessary and sufficient conditions for the sum of the $X_k$ normalized by $b_n$ to satisfy an MDP.

**Theorem 2.2.** *The following are equivalent:*

1. *the random variables $(X_i)$ satisfy $\mathbb{E}X_1 = 0$ and*

$$\limsup_{n\to\infty} \frac{n}{b_n^2} \log\left[n\mathbb{P}\left(|X_1| > b_n\right)\right] = -\infty; \tag{2.3}$$

2. *the sequence $\frac{1}{b_n}\sum_{i=1}^n X_i$ satisfies a moderate deviation principle with good rate function $I$ such that $I(x) > 0$ for all $x \neq 0$ and*

$$\lim_{x\to\infty} I(x) = \lim_{x\to-\infty} I(x) = \infty. \tag{2.4}$$

**Lemma 2.5.** *If the random variable $X_1$ satisfies condition (2.3), then $\mathbb{E}X_1^2 < \infty$.*

**Remark 2.6.** With Lemma 2.5 it follows from the proof of Theorem 2.2 to be given below that the rate function in Theorem 2.2 is

$$I(t) = \frac{1}{2\mathbb{E}X_1^2}t^2. \tag{2.7}$$

*Proof of Lemma 2.5.* To see, that $\mathbb{E}X_1^2 < \infty$ we need to show that

$$\int_0^\infty xP(|X_1| > x)\mathrm{d}x < \infty.$$

Without loss of generality we choose $b_0 = 0$ and observe that since $b_n/n \downarrow 0$ we have that $b_{n+1}/b_n \le (n+1)/n$ and thus $b_{n+1} \le 2\,b_n$. We obtain

$$\int_0^\infty xP(|X_1| > x)\mathrm{d}x = \sum_{n=0}^\infty \int_{b_n}^{b_{n+1}} xP(|X_1| > x)\mathrm{d}x \le \sum_{n=0}^\infty \int_{b_n}^{b_{n+1}} b_{n+1}P(|X_1| > b_n)\mathrm{d}x$$

$$\le 2\sum_{n=0}^\infty \int_{b_n}^{b_{n+1}} \frac{b_n}{n}nP(|X_1| > b_n)\mathrm{d}x.$$

From (2.3) we get that there is $N_0 \in \mathbb{N}$ such that for all $n \ge N_0$ we have $nP(|X_1| > b_n) \le \mathrm{e}^{-b_n^2/n}$. From here we get

$$\int_0^\infty xP(|X_1| > x)\mathrm{d}x \le 2\sum_{n=0}^{N_0-1} \int_{b_n}^{b_{n+1}} \frac{b_n}{n}nP(|X_1| > b_n)\mathrm{d}x + 2\sum_{n=N_0}^\infty \int_{b_n}^{b_{n+1}} \frac{b_n}{n}\mathrm{e}^{-\frac{b_n^2}{n}}\mathrm{d}x$$

$$= 2\sum_{n=0}^{N_0-1} \int_{b_n}^{b_{n+1}} \frac{b_n}{n}nP(|X_1| > b_n)\mathrm{d}x + 2\sum_{n=N_0}^\infty \frac{1}{\sqrt{n}}(b_{n+1} - b_n)\frac{b_n}{\sqrt{n}}\mathrm{e}^{-\frac{b_n^2}{n}}.$$

Since

$$2\sum_{n=N_0}^\infty \frac{1}{\sqrt{n}}(b_{n+1} - b_n)\frac{b_n}{\sqrt{n}}\mathrm{e}^{-\frac{b_n^2}{n}} \le 2\sum_{n=N_0}^\infty \frac{1}{\sqrt{n}}(b_n - b_{n-1})\frac{b_n}{\sqrt{n}}\mathrm{e}^{-\frac{b_n^2}{n}}$$

$$\le 2\int_0^\infty x\mathrm{e}^{-x^2}\mathrm{d}x < \infty,$$

the proof is finished. $\qquad\square$

**Remark 2.8.** Ledoux [9] discusses the case of random variables with values in a real separable Banach space $B$. We only comment on the main additional difficulties in this case and skip details. In the case of Banach space valued random variables one important difference is that (2.3) no more implies the square integrability which now has to be assumed.

Hence we conjecture that the statement of our result in the case of Banach space valued random variables should read as follows:

Suppose that $\mathbb{E}f(X_1) = 0$ and $\mathbb{E}f^2(X_1) < \infty$ for every $f$ in the dual $B'$ of $B$.

Suppose moreover that the family $\{f^2(X_1); f \in B', \|f\| \leq 1\}$ is uniformly integrable (realized for example when $\mathbb{E}\|X_1\|^2 < \infty$), then the following are equivalent:

- the sequence $(1/b_n) \sum_{i=1}^{n} X_i$ converges to 0 in probability and condition (2.3) is fulfilled;
- the sequence $(1/b_n) \sum_{i=1}^{n} X_i$ satisfies a moderate deviation principle with a good rate function $I$.

The form of the rate function in this case is described in detail in [9]. Note that it looks much more involved than in our case. Also note that Ledoux's proof of the MDP upper bound in [9] – other than ours – is based on isoperimetric inequalities. For details concerning Banach space valued random variables see [10].

*Proof of Theorem 2.2* $1 \Rightarrow 2$. From Lemma 2.5 we get $\mathbb{E}X_1^2 < \infty$.

To prove the moderate deviation estimates we employ truncation and the Gärtner–Ellis theorem. For $\tau > 0$ and $j = 1, 2, \ldots$ define

$$X_j^\tau := X_j^\tau(n) := X_j \mathbb{I}_{\{|X_j| \leq \tau \frac{n}{b_n}\}}.$$

By the Gärtner–Ellis theorem (Th. 2.3.6. in [5]) the sequence $\frac{1}{b_n} \sum_{j=1}^{n} X_j^\tau$ satisfies an MDP. Indeed for $t \in \mathbb{R}$:

$$\frac{n}{b_n^2} \log \mathbb{E} \exp \left( t \frac{b_n}{n} \sum_{j=1}^{n} X_j^\tau \right) = \frac{n^2}{b_n^2} \log \mathbb{E} \exp \left( t \frac{b_n}{n} X_1^\tau \right) = \frac{n^2}{b_n^2} \log \left( 1 + \frac{t^2}{2} \frac{b_n^2}{n^2} \mathbb{E}(X_1^\tau)^2 + \mathcal{O} \left( t^3 \frac{b_n^3}{n^3} \mathbb{E}(X_1^\tau)^3 \right) \right),$$

where we used that $\mathbb{E}X_1 = 0$ and that $\frac{b_n}{n} X_1^\tau \leq \tau$. For each positive $\varepsilon > 0$ one has

$$|t^3| \frac{b_n}{n} |\mathbb{E}(X_1^\tau)^3| = |t^3| \frac{b_n}{n} \left| \mathbb{E} \left[ (X_1^\tau)^3 I_{\{|X_1| \leq \varepsilon \frac{n}{b_n}\}} \right] \right| + |t^3| \frac{b_n}{n} \left| \mathbb{E} \left[ (X_1^\tau)^3 I_{\{|X_1| > \varepsilon \frac{n}{b_n}\}} \right] \right|$$

$$\leq |t^3| \varepsilon \mathbb{E}(X_1^\tau)^2 + |t^3| \tau^3 \frac{n^2}{b_n^2} \mathbb{P} \left( |X_1| > \varepsilon \frac{n}{b_n} \right).$$

Now since $\mathbb{E}X_1^2 = 2 \int x P(|X_1| > x) \mathrm{d}x < \infty$ it follows that $x^2 P(|X_1| > x) \to 0$ as $x \to \infty$, which readily implies that

$$|t^3| \tau^3 \frac{n^2}{b_n^2} \mathbb{P} \left( |X_1| > \varepsilon \frac{n}{b_n} \right) = \frac{|t^3| \tau^3}{\varepsilon^2} \left( \varepsilon^2 \frac{n^2}{b_n^2} \right) \mathbb{P} \left( |X_1| > \varepsilon \frac{n}{b_n} \right) \to 0$$

as $n \to \infty$ for all $\varepsilon > 0$, $t$, and $\tau > 0$. Hence

$$\lim_{n \to \infty} |t^3| \frac{b_n}{n} \mathbb{E}(X_1^\tau)^3 \leq |t^3| \varepsilon \mathbb{E}(X_1^\tau)^2.$$

Since this is true for all $\varepsilon > 0$ we obtain $\lim_{n \to \infty} |t^3| \frac{b_n}{n} \mathbb{E}(X_1^\tau)^3 = 0$ and therefore we get for each $\tau > 0$

$$\lim_{n \to \infty} \frac{n}{b_n^2} \log \mathbb{E} \exp \left( t \frac{b_n}{n} \sum_{j=1}^{n} X_j^\tau \right) = \lim_{n \to \infty} \frac{n^2}{b_n^2} \log \left( 1 + \frac{t^2}{2} \frac{b_n^2}{n^2} \mathbb{E}(X_1^\tau)^2 + \mathcal{O} \left( |t^3| \frac{b_n^3}{n^3} \mathbb{E}(X_1^\tau)^3 \right) \right)$$

$$= \frac{t^2}{2} \lim_{n \to \infty} \mathbb{E}(X_1^\tau)^2 + \lim_{n \to \infty} \mathcal{O} \left( |t^3| \frac{b_n}{n} \mathbb{E}(X_1^\tau)^3 \right) = \frac{t^2}{2} \mathbb{E}(X_1)^2.$$

So by the Gärtner–Ellis theorem each of the sequences $\left( \frac{1}{b_n} \sum_{i=1}^{n} X_i^\tau \right)_n$ satisfies an MDP with speed $\frac{b_n^2}{n}$ and rate function

$$I(t) = \sup_x \left[ xt - \frac{x^2}{2} \mathbb{E}(X_1)^2 \right] = \frac{1}{2\mathbb{E}X_1^2} t^2.$$

Remark that the rate function does not depend on $\tau$.

Thus it remains to show that the sequences $\left(\frac{1}{b_n}\sum_{i=1}^n X_i^\tau\right)_n$ and $\left(\frac{1}{b_n}\sum_{i=1}^n X_i\right)_n$ are exponentially equivalent (see [5], Sect. 4.2.2). Now for each $\tau > 0, \delta > 0$

$$\lim_{n\to\infty}\frac{n}{b_n^2}\log\mathbb{P}\left(\frac{1}{b_n}\left|\sum_{i=1}^n X_i - X_i^\tau\right| > 2\delta\right)$$

$$\leq \lim_{n\to\infty}\frac{n}{b_n^2}\log\max\left[\mathbb{P}\left(\frac{1}{b_n}\left|\sum_{i=1}^n X_i - Y_i\right| > \delta\right),\mathbb{P}\left(\frac{1}{b_n}\left|\sum_{i=1}^n Y_i - X_i^\tau\right| > \delta\right)\right], \quad (2.9)$$

where $Y_i := Y_i(n) := X_i\mathbb{I}\{|X_i| \leq b_n\}$. By Assumption (2.3)

$$\lim_{n\to\infty}\frac{n}{b_n^2}\log\mathbb{P}\left(\frac{1}{b_n}\left|\sum_{i=1}^n X_i - Y_i\right| > \delta\right) \leq \lim_{n\to\infty}\frac{n}{b_n^2}\log\mathbb{P}\left(\exists i : |X_i| > b_n\right)$$

$$\leq \lim_{n\to\infty}\frac{n}{b_n^2}\log\left[n\mathbb{P}\left(|X_1| > b_n\right)\right] = -\infty. \quad (2.10)$$

On the other hand by the exponential Chebyshev inequality we obtain for each $t > 0, \delta > 0, \tau > 0, M > 0$ and $Z_1 := Y_1 - X_1^\tau$:

$$\frac{n}{b_n^2}\log\mathbb{P}\left(\frac{1}{b_n}\left|\sum_{i=1}^n Y_i - X_i^\tau\right| > \delta\right) \leq -t\delta + \frac{n}{b_n^2}\log\mathbb{E}\left(\exp\left(t\frac{b_n}{n}\sum_{i=1}^n |Z_i|\right)\right)$$

$$= -t\delta + \frac{n^2}{b_n^2}\log\mathbb{E}\left(\exp\left(t\frac{b_n}{n}|Z_1|\right)\right)$$

$$\leq -t\delta + \frac{n^2}{b_n^2}\left(\mathbb{E}\left(\exp\left(t\frac{b_n}{n}|Z_1|\right)\right) - 1\right) \quad (2.11)$$

$$\leq -t\delta + \frac{n^2}{b_n^2}(\mathrm{e}^M - 1)\mathbb{P}\left(|X_1| \geq \tau\frac{n}{b_n}\right) + \frac{n^2}{b_n^2}\int_{\mathrm{e}^M}^{\mathrm{e}^{b_n^2}}\mathbb{P}\left(\mathrm{e}^{t\frac{b_n}{n}|X_1|} \geq s\right)\mathrm{d}s.$$

In the last step we used that $|Z_1| \leq |X_1|$. By Chebyshev's inequality

$$\frac{n^2}{b_n^2}(\mathrm{e}^M - 1)\mathbb{P}\left(|X_1| \geq \tau\frac{n}{b_n}\right) \leq (\mathrm{e}^M - 1)E(X_1^2)/\tau^2 \quad (2.12)$$

which converges to zero as $\tau \to \infty$.

On the other hand by a change of variable

$$\frac{n^2}{b_n^2}\int_{\mathrm{e}^M}^{\mathrm{e}^{b_n^2}}\mathbb{P}\left(\mathrm{e}^{t\frac{b_n}{n}|X_1|} \geq s\right)\mathrm{d}s = \frac{n^2}{b_n^2}\int_M^{b_n^2}\mathbb{P}\left(|X_1| \geq \frac{sn}{tb_n}\right)\mathrm{e}^s\mathrm{d}s$$

and since $\frac{n}{b_n} \to \infty$ there is a subsequence $m_n = m(n, s, t)$ such that for $s \geq 1$

$$b_{m_n} \leq \frac{sn}{tb_n} \leq b_{m_n+1}$$

and $m_n = m(n, s, t) \to \infty$ as $n \to \infty$ for every fixed choice of $s, t$ with $s \geq 1$. Hence we can compute

$$\frac{n^2}{b_n^2}\mathbb{P}\left(|X_1| \geq \frac{sn}{tb_n}\right)\mathrm{e}^s \leq \frac{t^2}{s^2}\frac{b_{m_n+1}^2}{m_n}m_n\mathbb{P}\left(|X_1| \geq b_{m_n}\right)\mathrm{e}^s \leq t^2\frac{b_{m_n+1}^2}{m_n}\mathrm{e}^{-K_0\frac{b_{m_n}^2}{m_n}+s} \quad (2.13)$$

for all constants $K_0 > 0$ and all $m_n$ large enough (depending on $K_0$). The last inequality follows directly from the assumption that $s \geq 1$ and assumption (2.3). Observe that since $b_{m_n}/m_n \to 0$ we have that $b_{m_n+1}/b_{m_n} \leq m_{n+1}/m_n$ on the one hand. On the other hand the definition of $m_n$ implies that $m_n = \mathcal{O}(n)$ (since $b_n^2/n \to \infty$) hence that there is a constant $K_1 > 0$ such that $\frac{m_{n+1}}{m_n} \leq K_1$ and $\frac{b_n m_n}{b_{m_n} n} \leq K_1$. Moreover

$$s \leq \frac{t b_{m_n+1} b_n}{n} \leq K_1 \frac{t b_{m_n} b_n}{n} \leq K_1^2 \frac{t b_{m_n}^2}{m_n} \tag{2.14}$$

hence

$$\frac{b_{m_n}^2}{m_n} \geq \frac{s}{t K_1^2} \tag{2.15}$$

and

$$s \leq \frac{b_{m_n}^2}{m_n} t K_1^2. \tag{2.16}$$

Now set $m := m_n$, then (2.13–2.15) and (2.16) together yield

$$\frac{n^2}{b_n^2} \mathbb{P}\left(|X_1| \geq \frac{sn}{t b_n}\right) e^s \leq K_1^2 t^2 \frac{b_m^2}{m} e^{\frac{1}{2}\left(-K_0 \frac{b_m^2}{m} + K_1^2 \frac{t b_m^2}{m}\right)} e^{\frac{s}{2}\left(1 - \frac{K_0}{t K_1^2}\right)}$$

for all $m$ large enough and all $K_0 > 0$. Thus

$$\limsup_{n \to \infty} \frac{n^2}{b_n^2} \int_M^{b_n^2} \mathbb{P}\left(|X_1| \geq \frac{sn}{t b_n}\right) e^s \mathrm{d}s \leq \limsup_{n \to \infty} K_1^2 t^2 \frac{b_m^2}{m} e^{\frac{1}{2}\left(-K_0 \frac{b_m^2}{m} + K_1^2 \frac{t b_m^2}{m}\right)} \int_M^{b_n^2} e^{\frac{s}{2}\left(1 - \frac{K_0}{t K_1^2}\right)} \mathrm{d}s. \tag{2.17}$$

Since $K_0 > 0$ was arbitrary, the first term on the right and side of (2.17) converges to zero as $n \to \infty$ – and therefore also as $m \to \infty$ – and the integral is bounded. Thus

$$\limsup_{n \to \infty} \frac{n^2}{b_n^2} \int_M^{b_n^2} \mathbb{P}\left(|X_1| \geq \frac{sn}{t b_n}\right) e^s \mathrm{d}s \to 0 \tag{2.18}$$

as $n \to \infty$.

Putting this together we obtain from (2.9) with the help of (2.10) on the one hand and (2.11) on the other (using (2.12) and (2.18)) that

$$\lim_{\tau \to \infty} \lim_{n \to \infty} \frac{n}{b_n^2} \log \mathbb{P}\left(\frac{1}{b_n}\left|\sum_{i=1}^n X_i - X_i^\tau\right| > \delta\right)$$

$$\leq \lim_{\tau \to \infty} \lim_{n \to \infty} \frac{n}{b_n^2} \log \max\left[\mathbb{P}\left(\frac{1}{b_n}\left|\sum_{i=1}^n X_i - Y_i\right| > \delta\right), \mathbb{P}\left(\frac{1}{b_n}\left|\sum_{i=1}^n Y_i - X_i^\tau\right| > \delta\right)\right]$$

$$= -t\delta + \lim_{\tau \to \infty} \lim_{n \to \infty} \frac{n^2}{b_n^2}(e^M - 1)\mathbb{P}\left(|X_1| \geq \tau \frac{n}{b_n}\right) + \lim_{\tau \to \infty} \lim_{n \to \infty} \frac{n^2}{b_n^2} \int_{e^M}^{e^{b_n^2}} \mathbb{P}(e^{t\frac{b_n}{n}|X_1|} \geq s)\mathrm{d}s$$

$$= -t\delta.$$

Since this is true for all $t > 0$ and $\delta > 0$ we arrive at

$$\lim_{\tau \to \infty} \lim_{n \to \infty} \frac{n}{b_n^2} \log \mathbb{P}\left(\frac{1}{b_n}\left|\sum_{i=1}^n X_i - X_i^\tau\right| > \delta\right) = -\infty. \tag{2.19}$$

So by the considerations at the beginning of the proof, the exponential equivalence shown in (2.19) proves the assertion.

$2 \Rightarrow 1$: The main difficulty is to pass from the moderate deviations for sums of random variables to assertions over the large deviations for a single random variable. Here Lemma 2.6 in [10] is of use. It allows to relate the large deviations for sums of independent random variables to assertions over the large deviations of their maximum. It states that for a sequence $(Z_i)$ of independent, positive random variables, every $N \in \mathbb{N}$, and every $t > 0$ such that $\mathbb{P}(\max_{i \leq N} Z_i > t) \leq \frac{1}{2}$ the following inequality holds true:

$$\sum_{i=1}^{N} \mathbb{P}(Z_i > t) \leq 2\mathbb{P}\left(\max_{i \leq N} Z_i > t\right). \tag{2.20}$$

To pass from the maximum to a single random variable Lévy's inequality (Prop. 2.3 in [10]) is useful. Even though Lévy's inequality has been shown to hold for non-symmetric random variables in [8] as well, we apply the version of symmetric random variables to obtain (2.21) below. This is: if $(X_i)$ is a sequence of real-valued, symmetric random variables and for $k \in \mathbb{N}$ let $S_k := \sum_{i=1}^{k} X_i$, then for every integer $N$ and every $t > 0$

$$\mathbb{P}\left(\max_{i \leq N} |X_i| > t\right) \leq 2\,\mathbb{P}\big(|S_N| > t\big). \tag{2.21}$$

Hence, following the ideas already indicated in [10] and [2], we symmetrize the $X_i$: introduce an i.i.d. sequence $(X_i')_{i \in \mathbb{N}}$ which is also independent of $(X_i)_{i \in \mathbb{N}}$ and such that $\mathcal{L}(X_1) = \mathcal{L}(X_1')$. Then the sequences $\frac{1}{b_n} \sum_{i=1}^{n} X_i$ and $\frac{1}{b_n} \sum_{i=1}^{n} -X_i'$ obey a moderate deviations principle with speed $b_n^2/n$ and rate function $I(t) = I'(t)$. Hence also $\frac{1}{b_n} \sum_{i=1}^{n} (X_i - X_i')$ obeys a moderate deviations principle with speed $b_n^2/n$ and rate function $J(z) = \inf_{(x,y):x-y=z} I(x) + I'(y)$. (Note that, if we knew Cond. 2.3 already, the rate function $I$ could only be $I(t) = \frac{1}{2\mathbb{E}X_1^2}t^2$ and thus we would obtain $J(z) = \frac{1}{4\mathbb{E}X_1^2}z^2$. However, Cond. 2.3 is just what we want to show here.) Note that $J(\cdot)$ inherits property (2.4) from the rate function $I$. Indeed it holds:

$$\lim_{x \to \infty} J(x) = \lim_{x \to -\infty} J(x) = \infty. \tag{2.22}$$

Now for all $M > 0$:

$$\mathbb{P}(|X_1| \geq 2Mb_n)\mathbb{P}(|X_1'| < Mb_n) \leq \mathbb{P}(|X_1 - X_1'| \geq Mb_n) \tag{2.23}$$

on the one hand. On the other hand by the large deviations upper bound for $\frac{1}{b_n} \sum_{i=1}^{n} (X_i - X_i')$

$$\mathbb{P}\left(\left|\frac{1}{b_n}\sum_{i=1}^{n}(X_i - X_i')\right| \geq M\right) \leq 2\exp\left(-\frac{b_n^2}{n}\frac{1}{2}J(M)\right). \tag{2.24}$$

Since the $(X_i - X_i')$ are symmetric, equation (2.21) together with (2.24) yields

$$\mathbb{P}\left(\max_{i=1,\dots,n}\left|\frac{1}{b_n}(X_i - X_i')\right| \geq M\right) \leq 4\exp\left(-\frac{b_n^2}{n}\frac{1}{2}J(M)\right). \tag{2.25}$$

By Lemma 2.6 in [10], *i.e.* by (2.20), if $n$ is so large that $4\exp\left(-\frac{b_n^2}{n}\frac{1}{2}J(M)\right) \leq 1/2$ we obtain from (2.25)

$$n\mathbb{P}\left(\left|\frac{1}{b_n}(X_i - X_i')\right| \geq M\right) \leq 2\mathbb{P}\left(\max_{i=1,\dots,n}\left|\frac{1}{b_n}(X_i - X_i')\right| \geq M\right)$$

$$\leq 8\exp\left(-\frac{b_n^2}{n}\frac{1}{8}J(M)\right). \tag{2.26}$$

Equations (2.26) and (2.23) together yield for each $M > 0$

$$\limsup_{n\to\infty} \frac{n}{b_n^2} \log\left(n\,\mathbb{P}(|X_1| \ge 2Mb_n)\right) = \limsup_{n\to\infty} \frac{n}{b_n^2} \log\left(n\,\mathbb{P}(|X_1| \ge 2Mb_n)\mathbb{P}(|X_1'| < Mb_n)\right)$$

$$\le \limsup_{n\to\infty} \frac{n}{b_n^2} \log\left(n\,\mathbb{P}(|X_1 - X_1'| \ge Mb_n)\right) \le -\frac{1}{2}J(M). \tag{2.27}$$

This at first glance looks rather unfortunate since simply substituting $M = 1/2$ does not yield the desired result. The point will be to exploit the quadratic character of the rate function as follows: for $M > 1$ we choose $m = m(n, M)$ such that

$$2Mb_m \le b_n < 2Mb_{m+1}. \tag{2.28}$$

Observe that since $b_n \ll n$ we know that $\liminf_{n\to\infty} \frac{n}{m} \ge 2M$ and since $b_n \gg \sqrt{n}$ we also know that $\limsup_{n\to\infty} \frac{n}{m} \le 4M^2$. This implies for all $M > 1$

$$\limsup_{n\to\infty} \frac{n}{b_n^2} \log\left(n\,\mathbb{P}(|X_1| \ge b_n)\right) \le \limsup_{n\to\infty} \frac{n}{b_n^2} \log\left(n\,\mathbb{P}(|X_1| \ge 2Mb_m)\right) \tag{2.29}$$

$$\le \limsup_{n\to\infty} \frac{m}{b_m^2}\frac{n\,b_m^2}{m\,b_n^2} \log\left(4M^2 m\,\mathbb{P}(|X_1| \ge 2Mb_m)\right)$$

$$\le \limsup_{m\to\infty} \frac{m}{b_m^2} \log\left(4M^2 m\,\mathbb{P}(|X_1| \ge 2Mb_m)\right)$$

$$\le -\frac{1}{2}J(M), \tag{2.30}$$

using (2.27) for the last inequality. Since this is true for all $M > 1$ we obtain from (2.22) the desired (2.3):

$$\limsup_{n\to\infty} \frac{n}{b_n^2} \log\left(n\,\mathbb{P}(|X_1| \ge b_n)\right) = -\infty.$$

Moreover, if $\mathbb{E}X_1 \ne 0$ by the weak law of large numbers $\frac{1}{b_n}\sum_{i=1}^{n}(X_i - \mathbb{E}X_i)$ converges to 0 weakly. Hence for any neighborhood $G$ of $\mathbb{E}X_1$ with $\overline{G} \not\ni 0$

$$\mathbb{P}\left(\frac{1}{b_n}\sum_{i=1}^{n} X_i \in G\right) \ge \frac{1}{2} \tag{2.31}$$

for $n$ large enough in contradiction to the moderate deviations upper bound. □

Theorem 2.2 gives a precise characterization of those distributions permitting a MDP for $\frac{1}{b_n}\sum_{i=1}^{n} X_i$ in terms of their probabilities and the sequence $(b_n)$. However, one might wonder which distributions satisfy condition (2.3). This is discussed in the following remark.

**Remark 2.32.** If the random variable $X_1$ has a local exponential moment, *i.e.* if

$$\mathbb{E}e^{t|X_1|} < \infty \qquad \text{for some } t > 0 \tag{2.33}$$

then (2.3) is satisfied for every increasing sequence of numbers $(b_n)$ with

$$\frac{b_n}{\sqrt{n}} \uparrow \infty \qquad \text{and} \qquad \frac{b_n}{n} \downarrow 0. \tag{2.34}$$

Indeed applying the exponential Chebyshev inequality yields with $t$ chosen as in (2.33)

$$\lim_{n\to\infty} \frac{n}{b_n^2} \log\left[n\mathbb{P}\left(|X_1| > b_n\right)\right] \le \lim_{n\to\infty} \left(\frac{n\log n}{b_n^2} - t\frac{n}{b_n} + \log\frac{n}{b_n^2}\mathbb{E}\mathrm{e}^{t|X_1|}\right)$$

$$= \lim_{n\to\infty} \frac{n}{b_n}\left(\frac{\log n}{n} - t\right) = \lim_{n\to\infty} -t\frac{n}{b_n} = -\infty.$$

Here we used that (2.33) implies that $\lim_{n\to\infty} \log\frac{n}{b_n^2}\mathbb{E}\mathrm{e}^{t|X_1|} = 0$ and that moreover $0 \le \lim_{n\to\infty}\frac{\log n}{b_n}$ $\le \lim_{n\to\infty}\frac{\sqrt{n}}{b_n} = 0$.

As a matter of fact in a certain sense the borderline case seems to be a class of distributions occurring in extreme value statistics, the so-called Weibull distributions. If $X_1$ is Weibull distributed, then $X_1 > 0$ with probability one and

$$\mathbb{P}(|X_1| > x) = \mathbb{P}(X_1 > x) = \mathrm{e}^{-cx^\alpha}$$

for some $c > 0$ and $\alpha > 0$. Without loss of generality we choose $c = 1$. If the so-called shape parameter $\alpha \ge 1$, $X_1$ has a local exponential moment and we are back in the case discussed above.

If $\alpha < 1$, then

$$\lim_{n\to\infty} \frac{n}{b_n^2} \log\left[n\mathbb{P}\left(|X_1| > b_n\right)\right] = \lim_{n\to\infty} \frac{n}{b_n}\left(\frac{\log n}{b_n} - \frac{1}{b_n^{1-\alpha}}\right)$$

from which one can determine the ($\alpha$-dependent) scale from which on (2.3) breaks down. Note that – not surprisingly – this breakdown invades from the large deviations scale into the direction of the CLT.

Let us next state that the result of Theorem 2.2 is a typical moderate deviations result. The generalization to the regime of large deviations does not hold true. This is made precise in the following remark:

**Remark 2.35.** If $b_n = n$, *i.e.* in the regime of large deviations Theorem 2.2 does not hold true. Indeed, in this case there are i.i.d. random variables such that their average obeys a large deviations principle but (2.3) is not fulfilled.

For example, take a sequence $Y_1, Y_2, \ldots$ of i.i.d. standard Gaussians in $\mathbb{R}$, *i.e.* $Y_1 \sim \mathcal{N}(0,1)$. Put $X_i = Y_i^2$ for $i = 1, 2 \ldots$ Then – since $\Lambda(t) = \log E\mathrm{e}^{tX_1}$ is finite for $t < \frac{1}{2}$ – Theorem 2.2.3 in [5] applies and we see that the sequence $(\frac{1}{n}\sum_{i=1}^n X_i)_n$ satisfies a large deviations principle with speed $n$ and good rate function

$$\Lambda^*(t) = \sup_{\lambda\in\mathbb{R}} \left[\lambda t - \Lambda(t)\right]. \tag{2.36}$$

On the other hand the sequence $(X_i)$ does not satisfy assumption (2.3). Indeed,

$$\limsup_{n\to\infty} \frac{1}{n} \log\left[n\mathbb{P}\left(|X_1| > n\right)\right] = \limsup_{n\to\infty} \frac{1}{n} \log\left[\mathbb{P}\left(|X_1| > n\right)\right]$$

$$= \limsup_{n\to\infty} \frac{1}{n} \log\left[2\mathbb{P}\left(|Y_1| > \sqrt{n}\right)\right] \tag{2.37}$$

$$= \limsup_{n\to\infty} \frac{1}{n} \log\left[\frac{1}{\sqrt{2\pi}} \int_{\sqrt{n}}^{\infty} \mathrm{e}^{-x^2/2}\mathrm{d}x\right]$$

$$= \limsup_{n\to\infty} \frac{1}{n} \log \mathrm{e}^{-n/2}$$

$$= -1/2$$

where we have applied standard estimates for Gaussian random variables.

This article has partially been written during a Research In Pairs project in Oberwolfach. We are grateful to the Mathematisches Forschungsinstitut Oberwolfach for its kind hospitality. We are indebted to an unknown referee for bringing the references [9] and [6] to our attention.

## References

[1] A. de Acosta, Moderate deviations and associated Laplace approximations for sums of independent random vectors. *Trans. Amer. Math. Soc.* **329** (1992) 357-375.

[2] M.A. Arcones, *The large deviation principle for empirical processes.* Preprint (1999).

[3] M. van den Berg, E. Bolthausen and F. den Hollander, Moderate deviations for the volume of the Wiener sausage. *Ann. Math.* **153** (2001) 355-406.

[4] H. Cramér, *Sur un nouveau théorème-limite de la théorie des probabilités*, Actualités Scientifique et Industrielles (736 Colloque consacré à la théorie des probabilités). Hermann (1938) 5-23.

[5] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications.* Springer, New York (1998).

[6] M. Djellout, Moderate deviations for martingale differences and applications to $\phi$-mixing sequences. *Stochastics and Stochastic Reports* (to appear).

[7] P. Eichelsbacher and U. Schmock, *Rank-dependent moderate deviations for U-empirical measures in strong topologies* (submitted).

[8] E. Giné and V. de la Peña, *Decoupling: From dependence to independence.* Springer-Verlag (1999).

[9] M. Ledoux, Sur les déviations modérées des sommes de variables aléatoires vectorielles indépendantes de même loi. *Ann. Inst. H. Poincaré* **28** (1992) 267-280.

[10] M. Ledoux and M. Talagrand, *Probability in Banach Spaces.* Springer-Verlag, Berlin (1991).

[11] M. Löwe and F. Merkl, Moderate deviations for longest increasing subsequences: The upper tail. *Comm. Pure Appl. Math.* **54** (2001) 1488-1520.